Engineering E. coli Nissle 1917 to advance and facilitate its use in biomedical applications

HARVARD UNIVERSITY

Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Harvard John A. Paulson School of Engineering and Applied Sciences
have examined a dissertation entitled:

"Engineering E. coli Nissle 1917 to advance and facilitate its use in biomedical applications"

presented by: Ilia Gelfat

Signature _____
Typed name: Professor S. Mitragotri

Signature _____
Typed name: Professor N. Joshi

Signature _____
Typed name: Professor D. Mooney

Signature _____
Typed name: Professor J. Leong

May 3, 2021

Engineering *E. coli* Nissle 1917 to advance and facilitate its use in biomedical applications

A dissertation presented

by

Ilia Gelfat

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

May 2021

Dissertation advisors: Prof. Neel Joshi, Prof. Samir Mitragotri Ilia Gelfat

Engineering *E. coli* Nissle 1917 to advance and facilitate its use in biomedical applications

Abstract

The concept of engineered living therapeutics refers to the utilization and reprogramming of living cells for a wide range of medical applications. Currently, most examples focus on engineering microbes for use in the human gut, with many choosing the strain *E. coli* Nissle 1917 (EcN) as a chassis organism due to its safety profile, probiotic properties, and the availability of compatible genetic tools. The number of proposed applications continues to increase, supported by our growing understanding of the human microbiome and advances in the field of synthetic biology. In this dissertation, I describe the development of new tools to further facilitate EcN engineering, as well as propose a novel approach for the use of engineered EcN to combat infectious disease.

We engineered and characterized two cryptic plasmids endogenous to EcN – pMUT1 and pMUT2. By inserting and testing several heterologous functional and regulatory components, we demonstrated the pMUT plasmids can be used as synthetic biology vectors in EcN. The removal of native pMUTs required the adaptation and optimization of existing plasmid curing techniques, and in the process shed light on the importance of the toxin-antitoxin system in pMUT2 for plasmid retention. Notably, we show that the engineered pMUTs can be used to express proteins *in vivo* and are retained during transit through the gastrointestinal tract without administration of antibiotics.

Next, we utilized EcN to display a network of functionalized protein fibers designed to bind and sequester pathogens. Building on the previously developed BIND platform, we used curli fibers, the main proteinaceous components of *E. coli* biofilms, fused to nanobodies – small antibody fragments derived from camelids. We demonstrate binding to several enteric pathogens, including enteropathogenic *E. coli*, *Shigella flexneri* and *Cryptosporidium parvum*, as well as the potent bacterial toxin Stx2. The work also describes the generation and testing of new nanobodies targeting several antigens associated with pathogenic *E. coli* strains.

Overall, the work described herein advances EcN engineering by providing novel genetic tools for its use in research and in the clinic, as well as demonstrating novel therapeutic functionalities, contributing to the fields of synthetic biology and engineered living therapeutics.

Table of contents

Acknowledgements

The work described in this dissertation would not have been possible without the help and support of many people, which I have been very fortunate to have around me over the past five years. First and foremost, I would like to thank my advisor, Prof. Neel Joshi, for his guidance and mentorship. During my time in the Joshi Lab, Neel has always been there to provide scientific feedback and practical advice while also allowing me the freedom to pursue the projects and research directions that I found most exciting and relevant. Even as our laboratory was moving from the Wyss Institute to Northeastern, Neel has consistently made time to meet with me individually, to discuss not only the progress of my work, but also my professional and personal development both in the PhD program and beyond.

I am also deeply grateful to all past and present members of the Joshi Lab, that have made my time here all the more memorable and valuable. In particular, I would like to thank Anna Duraj-Thatte, for the scientific mentorship early on and the many interesting and entertaining conversations on any and all topics; Pichet (Bom) Praveschotinunt, for the help with and advice on everything from mouse experiments to the job search process; Daniel Birnbaum, for the discussions of experiment ideas, music recommendations and the occasional jam session; Avinash Manjula-Basavanna, for the extensive material science expertise and the interesting conversations that follow; Anton Kan, for the fruitful collaborations and the fun stories at Wyss happy hours; and many other wonderful people, including Giorgia Cannici, Sivaram Emani, Miguel Suástegui, Junling Guo, Noémie-Manuelle Dorval Courchesne, Eneko Axpe, Ilona den Hartog, Arjirios Sourlis, Franziska Bahl, Richie Tay, Frank Ureña, Andyna Vernet, and Peter Nguyen. I would also like to thank the undergraduate and graduate students I had the honor of mentoring over the

years: Amy Zhou, Sanah Langer, and Colter Giem – it has been a pleasure working with you, and I have no doubt you are destined for great things!

Next, I would like to thank my dissertation committee members, Prof. Samir Mitragotri, Prof. David Mooney, and Prof. John Leong, for the mentorship they have provided over the years in our committee meetings. Their feedback and expertise have been instrumental in keeping my research on track, even as the COVID-19 pandemic required nontrivial adjustments to my experimental plans.

In addition to his role on the dissertation committee, Prof. Leong has also been a close collaborator, along with Prof. Charles Shoemaker. Neel and I have been extremely fortunate to have Prof. Leong and Prof. Shoemaker as collaborators – their genuine enthusiasm and belief in our platform, combined with their extensive expertise and open communication style, has propelled our shared work to new heights. For all this and more, my sincere gratitude goes to them and to the many other people in their respective labs and institutions who have worked on our shared projects: Yousuf Aqeel, Jacqueline Tremblay, Justyna Jaskiewicz, Anishma Shrestha, James N. Lee, Shenglan Hu, Xi Qian, Loranne Magoun, Prof. Abhineet Sheoran, Prof. Daniela Bedenice, Marcia Osburne, and Prof. Saul Tzipori. I would also like to recognize Prof. Georg Gerber for his helpful insights and input on the development of the mathematical model described in the appendix.

Next, I would like to give a heartfelt thank you to all the friends that have made my time outside the lab so special. From weekend trips to impromptu house parties, from roaming the corridors of Perkins Hall to the shared meals at the sub-par cafeteria formerly known as “Dudley House”, it is these moments that I will remember most when looking back at my time at Harvard. It would be impossible to list all the wonderful people that have been a part of my life here in Cambridge, but since I am not aware of a page limit, I will give it a

shot anyway: thank you to Lauren Christopher, Sophie Gilmore, Vivian Aluoch, Alyson Spitzig, Michael Williams, Uttkarsh Narayan, Dimitris Kalimeris, Yamini Bansal, Michael Han, Austin Vanbastelaer, Mickey Pearlstein-Gluck, Christina Chang, David Zhang, John Ahrens, Alex Najibi, Tara Sowrirajan, Rishi Singh, Laith Alhussein, George Abraham, Yanish Tucker, Shankar Nair, Yasser Hafiz, Kalki Kukreja, Raja Gopalakrishnan, Weilu Shen, Chara Podimata, Lindsey Brown, Heer Joisher, Rashmi Singh, Gabriel Alberts, Taylor Valley, Sean Kim, Kwasi Adu-Berchie, and Michael Florea.

I would also like to thank all of the friends back home in Israel and elsewhere around the world that have managed to keep in touch despite the distance and have been there for me throughout these past years: Iliya Gertmen, Alex Khosid, Bar Gazit, Amir Sivan, Izik Golan, Erez Zinman, Itamar Getzler, Oded Goldfracht, Omri Blum, Uri Iger, David Ginsburg, Amit Zalman, Chen Ferrera, and the entire "Pagim" collective; as well as Omri Caspi, Oz Meir, Yana Segal, Eitan Pevzner, Emil Sweid, Rana Geries, and Ross Cohen.

An extra special thank you goes to Tanvi Ranjan, whose love and support have brought me some of my best times of the past few years, and have also carried me through some of the toughest. Tanvi's contagious energy and caring spirit have truly been a ray of sunshine in the harsh Boston winter.

Lastly, I would like to thank my family, who has been an incredible source of strength throughout this entire period. The past five years have not been easy, with the last one being harder than most, but it was your consistent belief in me that has allowed me to believe in myself as well. Even though we are continents apart, the comfort that you give works just as well from far away. Thank you.

Chapter 1

Introduction

1.1. Engineered living therapeutics

1.1.1. From probiotics to engineered microbes

The human microbiome – the collection of microorganisms residing inside, on the surface of, or in association with the human body – is increasingly recognized for its role in health and disease. In recent decades, our understanding of these complex ecosystems of bacteria, archaea, fungi, protists, and viruses has evolved: whereas once these microbes were thought of as passive bystanders, we now know the microbiome to be implicated in a wide variety of functions ranging from immune modulation to the production of neurotransmitters. The study of the human microbiome has involved not only the characterization of its composition in health and disease states, but also demonstrations of how interventions in the microbiome can alter its normal function or disease progression[1,2]. In particular, the gut microbiome has become the focus of much of this research, as the gastrointestinal tract is home to the highest abundance and diversity of microbial life in the human body[2].

In parallel with our growing knowledge of the microbiome, the use of living microbes to improve health has also increased in the last few decades, typically in the form of probiotics. Probiotics are live microorganisms that, when administered in sufficient amounts, confer a health benefit to the host. Probiotic strains are usually human commensals or microbes derived from fermented food products[3]. Although the efficacy of probiotics may vary between formulations, there is evidence in support of their use for specific indications, including restoration of normal microbiota following treatment with antibiotics[4]. More recently, researchers have begun to develop novel therapeutic strategies based on the administration of communities of live microbes, either cultured in the lab (i.e.,

designed microbial consortia) or directly derived from the stool of a healthy donor (i.e., fecal microbiome transplant)[5].

Alongside the promise of probiotics and related therapeutic strategies, many have identified immense potential in genetically engineering microbes to specifically target disease. Microbes naturally possess the ability to sense their environment and respond to changes within it. Combined with the expanding toolkit available for genetic engineering, this allows for the utilization of microbes as programmable and responsive diagnostic or therapeutic agents that can produce a desired output *in situ*, at a site where such microbes are naturally present. Furthermore, the ability of engineered microbes to self-regenerate and locally replenish their population could result in consistent, longer-lasting therapeutic effects, likely reducing the number of required doses as well as production costs. This concept is often referred to as "engineered living therapeutics" or "engineered probiotics" [6-8].

1.1.2. Development of engineered living therapeutics

Some of the earliest examples of engineered living therapeutics utilized lactic acid bacteria (LAB) as a chassis organism, due to their long history of industrial use in fermented foods and well-established safety profile[9]. Steidler and coworkers engineered *Lactococcus lactis* to produce and secrete the murine cytokines interleukin- (IL-) 2 [10], IL-6 [11], and IL-10 [12]. In the latter study, the IL-10-producing *L. lactis* strain was shown to reduce the severity of inflammation in a mouse model of inflammatory bowel disease (IBD). Similar results were also obtained by using *L. lactis* to secrete a class of anti-inflammatory peptides called trefoil factors (TFF)[13]. LAB have also been proposed as mucosal vaccine delivery systems. Mansour and coworkers used an engineered *Lactobacillus gasseri* strain to express an antigen associated with the pathogen *Streptococcus pyogenes*[14]. In addition to these immunomodulatory and anti-inflammatory functions, LAB strains have also been

engineered to combat metabolic disorders such as phenylketonuria[15], diabetes and hyperglycemia[16].

More recently, Bacteroides species have emerged as potential chassis organisms for engineered living therapeutic applications. Along with Firmicutes, Bacteroides are among the most abundant inhabitants of the human gut microbiome, with the two phyla accounting for approximately 80% of individual organisms[17]. Unlike LAB, which are typically rapidly cleared after administration, the use of Bacteroides opens the possibility of long-term colonization, which may be particularly well-suited for the treatment of chronic disease. To this end, Mimee and coworkers have developed a genetic toolkit to facilitate the engineering of the human commensal *Bacteroides thetaiotamicron* and demonstrated its functionality *in vivo*[18]. Farrar and coworkers were able to utilize an additional *Bacteroides* strain, *Bacteroides ovatus*, to produce and secrete murine IL-2 in response to a chemical inducer[19]. *B. thetaiotamicron* has also been engineered to produce outer membrane vesicles for vaccine applications, successfully packaging antigens from *Salmonella enterica* serovar Typhimurium and Influenza virus A[20].

While engineered LAB, Bacteroides species, and a handful of other microbial strains continue to be developed for a variety of applications, the most common class of chassis organism currently utilized is *Escherichia coli* Nissle 1917 (EcN)[6].

1.1.3. *E. coli* Nissle 1917 and its use in the clinic

EcN was first isolated by German physician Alfred Nissle in 1917 from the stool of a military officer exhibiting particularly good gastrointestinal health compared to his fellow soldiers during an outbreak of diarrheal disease. Following this discovery and subsequent experimentation, the strain began to be used as a probiotic and is still marketed as such in

Europe under the trade name "Mutaflor" [21,22]. EcN has also been shown to ameliorate IBD – a clinical trial demonstrated that EcN was as effective as mesalazine, the standard drug used to treat ulcerative colitis, for maintaining remission[23]. This has been attributed in part to the lipopolysaccharide composition of the EcN outer membrane, which has been shown to induce the production of epithelial defensin, and promote restoration of tight junctions between enterocytes[22].

Several additional studies have shed light on the specific mechanisms involved in the probiotic properties of EcN. It has long been known that EcN produce microcins – low-molecular weight peptides that exhibit bactericidal activity against related microbes[24,25]. Later studies identified two microcins produced by EcN – MccM and MccH47 [26,27]. It is clear, however, that the production of microcins is not the only mechanism involved in the protective capacity of EcN against various pathogens, as microcins were not required for the ability of EcN to inhibit host cell invasion by *Salmonella enterica* serovar Typhimurium, *Shigella flexneri*, *Yersinia enterocolitica* and *Listeria monocytogenes*[28]. Hancock and coworkers have demonstrated the ability of EcN to outcompete several pathogenic *E. coli* strains during biofilm formation, including enteropathogenic, enterotoxigenic, and enterohemorrhagic *E. coli* (EPEC, ETEC and EHEC, respectively). This may explain the prophylactic effect of EcN for some diarrheal diseases[29]. Subsequent research has also shown that EcN can outcompete the pathogenic strain *E. coli* O157:H7 by occupying a similar nutritional niche[30].

1.1.4. Engineered *E. coli* Nissle 1917

The extensive use of *E. coli* as a model organism for biological research has resulted in an abundance of techniques for *E. coli* culture, isolation, and genetic manipulation. This, combined with its long track record of safe use in humans as a probiotic strain, has led to

EcN becoming the chassis of choice for a growing number of engineered living therapeutic and diagnostic projects.

Building on the abilities of wild-type EcN to combat enteric infections and interact with the host immune system, several groups have engineered EcN to treat inflammatory and infectious diseases. Geldart and coworkers created an antimicrobial peptide-producing strain of EcN to treat vancomycin-resistant *Enterococci* infections[31]. A similar approach was taken by Palmer and coworkers, engineering EcN to produce the microcin H47 in response to tetrathionate – a molecule produced in the gut during *Salmonella* infection[32]. Sarate and coworkers used EcN to produce a birch and grass antigen chimera and demonstrated that its intranasal administration was able to prevent allergic polysensitization in mice[33]. Several approaches utilized EcN to treat IBD, including expression of IL-10 [34], display of TFF[35], and production of the nematode-derived immunomodulatory protein cystatin[36].

EcN has also been engineered to treat various metabolic diseases. Isabella and coworkers utilized EcN to treat phenylketonuria, a disease characterized by an inability to metabolize phenylalanine, by engineering the microbes to express phenylalanine degrading enzymes[37]. A similar approach was used by Kurtz and coworkers to design an ammonia degrading EcN strain for treatment of hyperammonemia[38]. In other work, Somabhai *et al.* created an EcN strain capable of metabolizing excess dietary fructose, and demonstrated its ability to reduce the severity of fructose-induced metabolic syndrome in rats[39]. EcN has also been engineered to secrete the proteins GLP-1 and PDX-1, both of which have been shown to stimulate insulin production in intestinal epithelial cells, for the treatment of diabetes[40].

The ability of EcN to target solid tumors and grow within them has also led to the development of numerous strains for cancer diagnostic and therapeutic applications[41].

Stritzker and coworkers utilized this tumor targeting property to show that intravenously injected EcN can not only replicate within tumors, but also turn on the expression of heterologous genes *in situ* in response to the administration of the chemical inducer L-arabinose[42]. Danino and coworkers used EcN as a diagnostic for liver metastases by engineering the strain to express the reporter gene *lacZ*, which can cleave a colorimetric substrate rendering it detectable in urine[43]. In this example, the presence of EcN was itself used as the diagnostic output, as systemically administered EcN cannot typically survive outside tumors in immunocompetent individuals[22]. Subsequent work demonstrated that EcN can be used to locally deliver antibody fragments for checkpoint blockade therapy[44]. The examples described here account for only a small part of engineered EcN development, and the full scope includes its use for vaccine delivery[45], production of omega-3 fatty acids[46], and numerous other applications[41,47,48].

1.1.5. Translation of engineered living therapeutics into the clinic

The advances in the field of engineered living therapeutics have led to founding of several companies dedicated to their clinical development and implementation. These include Synlogic, which utilizes EcN as a chassis organism; Novome, which employs Bacteroides; and ActoBiotics (acquired by Precigen in recent years), which uses engineered LAB. Though none have yet commercialized their products, some substantial progress has been made. In particular, Synlogic has successfully completed a Phase I clinical trial for the treatment of phenylketonuria[49] and has characterized the residence time of their engineered strains in humans[50].

However, despite the promise of engineered living therapeutics, challenges remain for successful clinical translation. For some applications, the barriers are primarily scientific: the complex interplay between microbe and host in health and disease is yet to be fully

elucidated, and a deeper understanding of the underlying mechanisms will be required to enable certain therapeutic avenues[51]. In other cases, however, the hurdles may be overcome through innovations in microbial engineering. This is arguably the case for the IL-10-producing *L. lactis* strain developed by ActoBiotics for the treatment of IBD: although the strain met the required safety standards in a Phase I clinical trial[52], it failed to demonstrate therapeutic efficacy[53], despite promising preclinical results[12]. Therefore, the development of novel tools and strategies for bacterial engineering constitutes an important goal for translational biomedical research.

1.2. Relevant components for engineered *E. coli* Nissle 1917

1.2.1. The synthetic biology toolbox

Efforts to develop engineered living therapeutics have been greatly facilitated by advances in synthetic biology – a field which endeavors to utilize living cells, biomolecules, and pathways as building blocks for novel applications. Conceptually, synthetic biologists view these biological elements as separable, engineerable components which, to a first approximation, work independently of each other. It is therefore the work of scientists and engineers in this field to identify, develop, and characterize such components, each gradually adding to the growing toolbox of synthetic biology[54].

Some of the seminal work in the field involved the design of toy models, in which bacterial cells were programmed to recapitulate the behavior of electrical components[55]. Gardner and coworkers were able to engineer *E. coli* cells to include a "toggle switch" – a gene-regulatory system featuring two stable states, switching from one to the other in response to a specific stimulus[56]. Following a similar approach, Elowitz and Leibler created the "repressilator" –

a cellular system that predictably oscillates between three states, based on a network of mutually repressing genetic components[57]. However, more recent work has increasingly shifted towards concrete biological applications. Indeed, much of the research into engineered living therapeutics and diagnostics referenced above falls under the umbrella of synthetic biology. The development of novel genetic components has too shifted *in vivo* in some cases. This is exemplified by the "memory circuit" designed by Kotula and coworkers – engineered *E. coli* cells containing this circuit were able to survive in the murine gastrointestinal tract, sense the presence of a chemical inducer, and respond by editing their own DNA, allowing for subsequent isolation and detection[58].

Notably, most synthetic biology projects rely on plasmids to introduce novel DNA sequences and express heterologous proteins. Compared to the incorporation of genetic material into the genome, the techniques involved in plasmid-based cellular engineering are substantially simpler, quicker, and more efficient. This is of particular importance in an engineering field such as synthetic biology, to speed up the iterative process of designing, building, and testing novel systems. Some of the research presented in this dissertation entails the development of novel plasmid systems for EcN engineering. Other biological components central to the work described herein include curli fibers and nanobodies and are discussed below in more detail.

1.2.2. Curli fibers

In *E. coli* and other Enterobacteriaceae species, curli fibers are the major proteinaceous component in biofilms. Along with polysaccharides, extracellular DNA, and several additional secreted proteins, curli fibers create a robust polymeric mesh that facilitates adhesion to surfaces and protects the bacteria embedded within it from unfavorable environmental conditions[59]. Curli belong to a class of structural proteins known as

functional amyloids – CsgA, the major curli subunit, adopts an amyloid fold characterized by a cross-beta structure, which endows the resulting fibers with resistance to chemical and proteolytic degradation[60-62].

In *E. coli*, curli fibers and their associated cellular machinery is encoded by the *csg* genes (i.e., curli specific genes), which are organized into a cluster of two divergent operons on the bacterial chromosome. The *csgBAC* operon encodes the CsgA monomer, as well as the minor curli subunit CsgB, which mediates fiber nucleation and attachment to the bacterial surface. CsgC serves as a periplasmic chaperone and prevents premature fiber formation inside the bacterium. The second operon, *csgDEFG*, contains the genes related to the dedicated secretion system associated with curli fibers. CsgG forms a nonameric pore in the *E. coli* outer membrane through which CsgA and CsgB are secreted into the extracellular space. Each CsgG subunit is associated with CsgE on the periplasmic side, which is involved in substrate recognition, and CsgF on its extracellular side, which mediates the surface attachment of fibers via its interaction with CsgB. The transcription of the *csg* operons is controlled by CsgD, which upregulates curli production in response to certain environmental signals such as low temperatures and nutrient-limited conditions. Following the secretion of unfolded CsgA, the monomers adopt the amyloid fold as they self-assemble into cell-anchored curli fibers[63-68].

Several properties of the curli system make it well-suited for use in synthetic biology applications. Unlike the extracellular polysaccharides and DNA that comprise biofilms, curli fibers are proteins, which makes it relatively straightforward to genetically fuse heterologous functional domains to them. By inserting a genetic sequence of interest downstream of the *csgA* gene, it is possible to achieve multivalent display of the desired component on the surface of an engineered bacterial chassis. The robustness of curli fibers

can also be advantageous when utilized in harsh environments such as the gastrointestinal tract. With these considerations in mind, our lab developed the BIND (biofilm integrated nanofiber display) system, demonstrating the ability of *E. coli* to express and secrete modified curli monomers that retain their capacity for self-assembly as well as the functionality of the fused heterologous domains[69].

Subsequent work elaborated on these capabilities and incorporated novel functionalities, including immobilization of enzymes on curli for biocatalysis applications[70], selective abstraction of rare earth metals using a curli-displayed binding domain[71], and binding of mercury for bioremediation[72]. Moser and coworkers appended peptide tags to CsgA, using the resulting curli fibers to template bacterial growth on a variety of surfaces, including plastics, ceramics, and textiles[73]. More recently, Pu and coworkers utilized an influenza binding peptide to endow curli fibers with the ability to remove virus particles from river water[74]. During my initial years in the Joshi Laboratory, I took part in our first efforts to utilize the BIND platform in a therapeutic context. By engineering EcN to produce CsgA-TFF fusions, we were able to demonstrate the ability of these functionalized fibers to reduce the severity and duration of experimentally induced colitis in mice[35]. Some of the work described herein builds upon these results, with the aim of continuing to advance therapeutic applications of engineered curli fibers.

1.2.3. Nanobodies – single domain antibody fragments

Antibodies are complex proteins produced by the immune system that can specifically bind to foreign entities. In addition to their native immune functions, antibodies play an integral role in modern biology and medicine, as they are used in various biological assays, diagnostics, and as therapeutic agents[75]. However, conventional antibodies cannot be used

in synthetic biology projects involving bacteria, since prokaryotes can typically only produce specifically engineered antibody fragments[76].

Camelid species such as camels, llamas and alpaca are known to produce antibodies that have a simpler structure – unlike conventional antibodies, where the antigen binding site occurs at the interface of two protein moieties known as the heavy chain and the light chain, these camelid antibodies feature only a heavy chain. The variable domain of these heavy chain-only antibodies, which contains the antigen binding site, can therefore retain its function even when separated from the remaining components. These domains, also known as VHHs or nanobodies, are small and simple enough to be expressed in bacteria and have subsequently found extensive use in synthetic biology[77].

Piñero-Lambea and coworkers were able to display nanobodies on the surface of *E. coli* cells, using them as synthetic adhesins that mediate attachment to antigen-coated surfaces as well as mammalian cells[78]. Lactic acid bacteria have been engineered to secrete a variety of VHHs for the treatment of inflammation and infectious disease, targeting tumor necrosis factor α[79], rotavirus[80], and the *Clostridium difficile* toxin TcdB[81], among other examples[82]. Similar approaches have been pursued in EcN[44,83]. Nanobodies have also been proposed as standalone therapeutics, targeting various pathogenic agents including EHEC[84], ETEC[85], and *Shigella flexneri*[86].

1.3. Dissertation overview

The work presented in this dissertation focuses on the development of EcN as a chassis organism for engineered living therapeutics. In keeping with the framework and goals of synthetic biology, I both describe the development of novel genetic components for use in EcN, as well as utilize existing components of the EcN engineering toolkit to implement new therapeutics for the treatment of infectious disease.

Chapter 2 reports the development and characterization of novel components for EcN engineering – the pMUT plasmids. While these plasmids are natively found in EcN, our work demonstrates they can be used as vectors for heterologous protein expression, both *in vitro* and *in vivo*. Unlike standard synthetic biology vectors, the engineered pMUT plasmids do not require antibiotic selection to be retained by EcN, opening the door to new uses in research and in the clinic.

In chapter 3, I describe a unique approach for treatment or prophylaxis of enteric infectious disease, which we have termed curli-based pathogen sequestration. By utilizing the BIND platform to express CsgA-nanobody fusions, we were able to create a robust mesh of functionalized curli fibers displayed on the surface of EcN. We demonstrate the ability of these curli-VHH-producing bacteria to bind to several enteric pathogens, including enteropathogenic *E. coli*, *Shigella flexneri* and *Cryptosporidium parvum*, as well as the potent bacterial toxin Stx2. Notably, this work also includes the generation and testing of novel VHHs against several pathogenic *E. coli* strains, and was done in close collaboration with the laboratories of Professor John Leong and Professor Charles Shoemaker. While several applications have been proposed for both curli fibers and nanobodies in the context

of treating infectious disease, their properties make them uniquely suited for the pathogen sequestration approach discussed here.

The conclusions of the work are presented in chapter 4, along with further discussion and potential directions for future research. Additional work is included in the appendix, where I describe the development of a mathematical model for the population dynamics of a probiotic species and an enteric pathogen as measured from fecal samples. Although the model was designed to help evaluate the viability of the pathogen sequestration approach from chapter 3 when applied *in vivo*, it may be of use for any experimental system that involves the coadministration of a pathogen and a probiotic strain.

1.4. References

1. Cho, I., & Blaser, M. J. (2012). The human microbiome: at the interface of health and disease. *Nature Reviews Genetics*, *13*(4), 260-270.
2. Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J. H., Chinwalla, A. T., ... & McCorrison, J. M. (2012). Structure, function and diversity of the healthy human microbiome. *nature*, *486*(7402), 207.
3. Williams, N. T. (2010). Probiotics. *American Journal of Health-System Pharmacy*, *67*(6), 449-458.
4. Hempel, S., Newberry, S. J., Maher, A. R., Wang, Z., Miles, J. N., Shanman, R., ... & Shekelle, P. G. (2012). Probiotics for the prevention and treatment of antibiotic-associated diarrhea: a systematic review and meta-analysis. *Jama*, *307*(18), 1959-1969.
5. Mimee, M., Citorik, R. J., & Lu, T. K. (2016). Microbiome therapeutics—advances and challenges. *Advanced drug delivery reviews*, *105*, 44-54.
6. Pedrolli, D. B., Ribeiro, N. V., Squizato, P. N., de Jesus, V. N., Cozetto, D. A., Tuma, R. B., ... & Cerri, M. O. (2019). Engineering microbial living therapeutics: the synthetic biology toolbox. *Trends in biotechnology*, *37*(1), 100-115.
7. Kelly, V. W., Liang, B. K., & Sirk, S. J. (2020). Living Therapeutics: The Next Frontier of Precision Medicine. *ACS synthetic biology*, *9*(12), 3184-3201.
8. Sola-Oladokun, B., Culligan, E. P., & Sleator, R. D. (2017). Engineered probiotics: applications and biological containment. *Annual review of food science and technology*, *8*, 353-370.
9. Mays, Z. J., & Nair, N. U. (2018). Synthetic biology in probiotic lactic acid bacteria: At the frontier of living therapeutics. *Current opinion in biotechnology*, *53*, 224-231.
10. Steidler, L., Wells, J. M., Raeymaekers, A., Vandekerckhove, J., Fiers, W., & Remaut, E. (1995). Secretion of biologically active murine interleukin-2 by Lactococcus lactis subsp. lactis. *Applied and environmental microbiology*, *61*(4), 1627-1629.
11. Steidler, L., Robinson, K., Chamberlain, L., Schofield, K. M., Remaut, E., Le Page, R. W., & Wells, J. M. (1998). Mucosal delivery of murine interleukin-2 (IL-2) and IL-6 by recombinant strains of Lactococcus lactis coexpressing antigen and cytokine. *Infection and immunity*, *66*(7), 3183-3189.
12. Steidler, L., Hans, W., Schotte, L., Neirynck, S., Obermeier, F., Falk, W., ... & Remaut, E. (2000). Treatment of murine colitis by Lactococcus lactis secreting interleukin-10. *Science*, *289*(5483), 1352-1355.
13. Vandenbroucke, K., Hans, W., Van Huysse, J., Neirynck, S., Demetter, P., Remaut, E., ... & Steidler, L. (2004). Active delivery of trefoil factors by genetically modified

Lactococcus lactis prevents and heals acute colitis in mice. *Gastroenterology*, *127*(2), 502-513.

14. Mansour, N. M., & Abdelaziz, S. A. (2016). Oral immunization of mice with engineered Lactobacillus gasseri NM713 strain expressing Streptococcus pyogenes M6 antigen. *Microbiology and immunology*, *60*(8), 527-532.

15. Durrer, K. E., Allen, M. S., & Hunt von Herbing, I. (2017). Genetically engineered probiotic for the treatment of phenylketonuria (PKU); assessment of a novel treatment in vitro and in the PAHenu2 mouse model of PKU. *PloS one*, *12*(5), e0176286.

16. Ma, Y., Liu, J., Hou, J., Dong, Y., Lu, Y., Jin, L., ... & Wu, J. (2014). Oral administration of recombinant Lactococcus lactis expressing HSP65 and tandemly repeated P277 reduces the incidence of type I diabetes in non-obese diabetic mice. *PloS one*, *9*(8), e105701.

17. Gosalbes, M. J., Durbán, A., Pignatelli, M., Abellan, J. J., Jiménez-Hernández, N., Pérez-Cobas, A. E., ... & Moya, A. (2011). Metatranscriptomic approach to analyze the functional human gut microbiota. *PloS one*, *6*(3), e17447.

18. Mimee, M., Tucker, A. C., Voigt, C. A., & Lu, T. K. (2015). Programming a human commensal bacterium, Bacteroides thetaiotaomicron, to sense and respond to stimuli in the murine gut microbiota. *Cell systems*, *1*(1), 62-71.

19. Farrar, M. D., Whitehead, T. R., Lan, J., Dilger, P., Thorpe, R., Holland, K. T., & Carding, S. R. (2005). Engineering of the gut commensal bacterium Bacteroides ovatus to produce and secrete biologically active murine interleukin-2 in response to xylan. *Journal of applied microbiology*, *98*(5), 1191-1197.

20. Carvalho, A. L., Fonseca, S., Miquel-Clopés, A., Cross, K., Kok, K. S., Wegmann, U., ... & Carding, S. R. (2019). Bioengineering commensal bacteria-derived outer membrane vesicles for delivery of biologics to the gastrointestinal and respiratory tract. *Journal of extracellular vesicles*, *8*(1), 1632100.

21. Schultz, M. (2008). Clinical use of E. coli Nissle 1917 in inflammatory bowel disease. *Inflammatory bowel diseases*, *14*(7), 1012-1018.

22. Sonnenborn, U., & Schulze, J. (2009). The non-pathogenic Escherichia coli strain Nissle 1917–features of a versatile probiotic. *Microbial Ecology in Health and Disease*, *21*(3-4), 122-158.

23. Kruis, W., Frič, P., Pokrotnieks, J., Lukáš, M., Fixa, B., Kaščák, M., ... & Schulze, J. (2004). Maintaining remission of ulcerative colitis with the probiotic Escherichia coli Nissle 1917 is as effective as with standard mesalazine. *Gut*, *53*(11), 1617-1623.

24. Papavassiliou, J. (1961). Biological characteristics of colicine X. *Nature*, *190*(4770), 110-110.

25. Sassone-Corsi, M., Nuccio, S. P., Liu, H., Hernandez, D., Vu, C. T., Takahashi, A. A., ... & Raffatellu, M. (2016). Microcins mediate competition among Enterobacteriaceae in the inflamed gut. *Nature*, *540*(7632), 280-283.

26. Patzer, S., Baquero, M. R., Bravo, D., Moreno, F., & Hantke, K. (2003). The colicin G, H and X determinants encode microcins M and H47, which might utilize the catecholate siderophore receptors FepA, Cir, Fiu and IroN. *Microbiology*, *149*(9), 2557-2570.

27. Vassiliadis, G., Destoumieux-Garzón, D., Lombard, C., Rebuffat, S., & Peduzzi, J. (2010). Isolation and characterization of two members of the siderophore-microcin family, microcins M and H47. *Antimicrobial agents and chemotherapy*, *54*(1), 288-297.

28. Altenhoefer, A., Oswald, S., Sonnenborn, U., Enders, C., Schulze, J., Hacker, J., & Oelschlaeger, T. A. (2004). The probiotic Escherichia coli strain Nissle 1917 interferes with invasion of human intestinal epithelial cells by different enteroinvasive bacterial pathogens. *FEMS Immunology & Medical Microbiology*, *40*(3), 223-229.

29. Hancock, V., Dahl, M., & Klemm, P. (2010). Probiotic Escherichia coli strain Nissle 1917 outcompetes intestinalpathogens during biofilm formation. *Journal of medical microbiology*, *59*(4), 392-399.

30. Maltby, R., Leatham-Jensen, M. P., Gibson, T., Cohen, P. S., & Conway, T. (2013). Nutritional basis for colonization resistance by human commensal Escherichia coli strains HS and Nissle 1917 against E. coli O157: H7 in the mouse intestine. *PloS one*, *8*(1), e53957.

31. Geldart, K. G., Kommineni, S., Forbes, M., Hayward, M., Dunny, G. M., Salzman, N. H., & Kaznessis, Y. N. (2018). Engineered E. coli Nissle 1917 for the reduction of vancomycin-resistant Enterococcus in the intestinal tract. *Bioengineering & translational medicine*, *3*(3), 197-208.

32. Palmer, J. D., Piattelli, E., McCormick, B. A., Silby, M. W., Brigham, C. J., & Bucci, V. (2018). Engineered probiotic for the inhibition of Salmonella via tetrathionate-induced production of microcin H47. *ACS infectious diseases*, *4*(1), 39-45.

33. Sarate, P. J., Heinl, S., Poiret, S., Drinić, M., Zwicker, C., Schabussova, I., ... & Wiedermann, U. (2019). E. coli Nissle 1917 is a safe mucosal delivery vector for a birch-grass pollen chimera to prevent allergic poly-sensitization. *Mucosal immunology*, *12*(1), 132-144.

34. Gardlik, R., Palffy, R., & Celec, P. (2012). Recombinant probiotic therapy in experimental colitis in mice. *Folia biologica*, *58*(6), 238.

35. Praveschotinunt, P., Duraj-Thatte, A. M., Gelfat, I., Bahl, F., Chou, D. B., & Joshi, N. S. (2019). Engineered E. coli Nissle 1917 for the delivery of matrix-tethered therapeutic domains to the gut. *Nature communications*, *10*(1), 1-14.

36. Whelan, R. A., Rausch, S., Ebner, F., Günzel, D., Richter, J. F., Hering, N. A., ... & Hartmann, S. (2014). A transgenic probiotic secreting a parasite immunomodulator for site-directed treatment of gut inflammation. *Molecular Therapy*, *22*(10), 1730-1740.

37. Isabella, V. M., Ha, B. N., Castillo, M. J., Lubkowicz, D. J., Rowe, S. E., Millet, Y. A., ... & Falb, D. (2018). Development of a synthetic live bacterial therapeutic for the human metabolic disease phenylketonuria. *Nature biotechnology*, *36*(9), 857-864.

38. Kurtz, C. B., Millet, Y. A., Puurunen, M. K., Perreault, M., Charbonneau, M. R., Isabella, V. M., ... & Miller, P. F. (2019). An engineered E. coli Nissle improves hyperammonemia and survival in mice and shows dose-dependent exposure in healthy humans. *Science translational medicine*, *11*(475).

39. Somabhai, C. A., Raghuvanshi, R., & Nareshkumar, G. (2016). Genetically engineered Escherichia coli Nissle 1917 synbiotics reduce metabolic effects induced by chronic consumption of dietary fructose. *PloS one*, *11*(10), e0164860.

40. Duan, F., Curtis, K. L., & March, J. C. (2008). Secretion of insulinotropic proteins by commensal bacteria: rewiring the gut to treat diabetes. *Applied and environmental microbiology*, *74*(23), 7437-7438.

41. Yu, X., Lin, C., Yu, J., Qi, Q., & Wang, Q. (2020). Bioengineered Escherichia coli Nissle 1917 for tumour-targeting therapy. *Microbial biotechnology*, *13*(3), 629-636.

42. Stritzker, J., Weibel, S., Hill, P. J., Oelschlaeger, T. A., Goebel, W., & Szalay, A. A. (2007). Tumor-specific colonization, tissue distribution, and gene induction by probiotic Escherichia coli Nissle 1917 in live mice. *International journal of medical microbiology*, *297*(3), 151-162.

43. Danino, T., Prindle, A., Kwong, G. A., Skalak, M., Li, H., Allen, K., ... & Bhatia, S. N. (2015). Programmable probiotics for detection of cancer in urine. *Science translational medicine*, *7*(289), 289ra84-289ra84.

44. Gurbatri, C. R., Lia, I., Vincent, R., Coker, C., Castro, S., Treuting, P. M., ... & Danino, T. (2020). Engineered probiotics for local tumor delivery of checkpoint blockade nanobodies. *Science translational medicine*, *12*(530).

45. Rosenthal, J. (2014). Engineered Outer Membrane Vesicles Derived From Probiotic Escherichia coli Nissle 1917 As Recobinant Subunit Antigen Carreirs For The Development Of Pathogen-Mimetic Vaccines.

46. Amiri-Jami, M., Abdelhamid, A. G., Hazaa, M., Kakuda, Y., & Griffths, M. W. (2015). Recombinant production of omega-3 fatty acids by probiotic Escherichia coli Nissle 1917. *FEMS microbiology letters*, *362*(20).

47. Ou, B., Yang, Y., Tham, W. L., Chen, L., Guo, J., & Zhu, G. (2016). Genetic engineering of probiotic Escherichia coli Nissle 1917 for clinical application. *Applied microbiology and biotechnology*, *100*(20), 8693-8699.

48. Aggarwal, N., Breedon, A. M. E., Davis, C. M., Hwang, I. Y., & Chang, M. W. (2020). Engineering probiotics for therapeutic applications: recent examples and translational outlook. *Current opinion in biotechnology*, *65*, 171-179.

49. National Library of Medicine (U.S.). (2018, April 17 – 2019, June 21). *Safety and Tolerability of SYNB1618 in Healthy Adult Volunteers and Adult Subjects With Phenylketonuria (PKU).* Identifier NCT03516487. https://clinicaltrials.gov/ct2/show/NCT03516487

50. Kurtz, C., Denney, W. S., Blankstein, L., Guilmain, S. E., Machinani, S., Kotula, J., ... & Brennan, A. M. (2018). Translational Development of Microbiome-Based Therapeutics: Kinetics of E. coli Nissle and Engineered Strains in Humans and Nonhuman Primates. *Clinical and translational science*, *11*(2), 200-207.

51. Tan, Y., Shen, J., Si, T., Ho, C. L., Li, Y., & Dai, L. (2020). Engineered Live Biotherapeutics: Progress and Challenges. *Biotechnology Journal*, *15*(10), 2000155.

52. Braat, H., Rottiers, P., Hommes, D. W., Huyghebaert, N., Remaut, E., Remon, J. P., ... & Steidler, L. (2006). A phase I trial with transgenic bacteria expressing interleukin-10 in Crohn's disease. *Clinical gastroenterology and hepatology*, *4*(6), 754-759.

53. Bron, P. A., & Kleerebezem, M. (2018). Lactic acid bacteria for delivery of endogenous or engineered therapeutic molecules. *Frontiers in microbiology*, *9*, 1821.

54. Cheng, A. A., & Lu, T. K. (2012). Synthetic biology: an emerging engineering discipline. *Annual review of biomedical engineering*, *14*, 155-178.

55. Perry, N., & Ninfa, A. J. (2012). Synthetic networks: oscillators and toggle switches for Escherichia coli. In *Synthetic Gene Networks* (pp. 287-300). Humana Press.

56. Gardner, T. S., Cantor, C. R., & Collins, J. J. (2000). Construction of a genetic toggle switch in Escherichia coli. *Nature*, *403*(6767), 339-342.

57. Elowitz, M. B., & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, *403*(6767), 335-338.

58. Kotula, J. W., Kerns, S. J., Shaket, L. A., Siraj, L., Collins, J. J., Way, J. C., & Silver, P. A. (2014). Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proceedings of the National Academy of Sciences*, *111*(13), 4838-4843.

59. Fong, J. N., & Yildiz, F. H. (2015). Biofilm matrix proteins. *Microbial Biofilms*, 201-222.

60. Erskine, E., MacPhee, C. E., & Stanley-Wall, N. R. (2018). Functional amyloid and other protein fibers in the biofilm matrix. *Journal of molecular biology*, *430*(20), 3642-3656.

61. Maury, C. P. J. (2009). The emerging concept of functional amyloid. *Journal of internal medicine*, *265*(3), 329-334.

62. Collinson, S. K., Emödy, L., Müller, K. H., & Kay, W. W. (1991). Purification and characterization of thin, aggregative fimbriae from Salmonella enteritidis. *Journal of bacteriology*, *173*(15), 4773-4781.

63. Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., ... & Hultgren, S. J. (2002). Role of Escherichia coli curli operons in directing amyloid fiber formation. *Science*, *295*(5556), 851-855.

64. Barnhart, M. M., & Chapman, M. R. (2006). Curli biogenesis and function. *Annu. Rev. Microbiol.*, *60*, 131-147.

65. Blanco, L. P., Evans, M. L., Smith, D. R., Badtke, M. P., & Chapman, M. R. (2012). Diversity, biogenesis and function of microbial amyloids. *Trends in microbiology*, *20*(2), 66-73.

66. Deshmukh, M., Evans, M. L., & Chapman, M. R. (2018). Amyloid by design: intrinsic regulation of microbial amyloid assembly. *Journal of molecular biology*, *430*(20), 3631-3641.

67. Goyal, P., Krasteva, P. V., Van Gerven, N., Gubellini, F., Van den Broeck, I., Troupiotis-Tsaïlaki, A., ... & Remaut, H. (2014). Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG. *Nature*, *516*(7530), 250-253.

68. Zhang, M., Shi, H., Zhang, X., Zhang, X., & Huang, Y. (2020). Cryo-EM structure of the nonameric CsgG-CsgF complex and its implications for controlling curli biogenesis in Enterobacteriaceae. *PLoS biology*, *18*(6), e3000748.

69. Nguyen, P. Q., Botyanszki, Z., Tay, P. K. R., & Joshi, N. S. (2014). Programmable biofilm-based materials from engineered curli nanofibres. *Nature communications*, *5*(1), 1-10.

70. Botyanszki, Z., Tay, P. K. R., Nguyen, P. Q., Nussbaumer, M. G., & Joshi, N. S. (2015). Engineered catalytic biofilms: Site-specific enzyme immobilization onto E. coli curli nanofibers. *Biotechnology and bioengineering*, *112*(10), 2016-2024.

71. Tay, P. K. R., Manjula-Basavanna, A., & Joshi, N. S. (2018). Repurposing bacterial extracellular matrix for selective and differential abstraction of rare earth elements. *Green Chemistry*, *20*(15), 3512-3520.

72. Tay, P. K. R., Nguyen, P. Q., & Joshi, N. S. (2017). A synthetic circuit for mercury bioremediation using self-assembling functional amyloids. *ACS synthetic biology*, *6*(10), 1841-1850.

73. Moser, F., Tham, E., González, L. M., Lu, T. K., & Voigt, C. A. (2019). Light-controlled, high-resolution patterning of living engineered bacteria onto textiles, ceramics, and plastic. *Advanced Functional Materials*, *29*(30), 1901788.

74. Pu, J., Liu, Y., Zhang, J., An, B., Li, Y., Wang, X., ... & Zhong, C. (2020). Virus disinfection from environmental water sources using living engineered biofilm materials. *Advanced Science*, *7*(14), 1903558.

75. Padlan, E. A. (1994). Anatomy of the antibody molecule. *Molecular immunology*, *31*(3), 169-217.

76. Fernández, L. A. (2004). Prokaryotic expression of antibodies and affibodies. *Current opinion in biotechnology*, *15*(4), 364-373.

77. Muyldermans, S. (2013). Nanobodies: natural single-domain antibodies. *Annual review of biochemistry*, *82*, 775-797.

78. Piñero-Lambea, C., Bodelón, G., Fernández-Periáñez, R., Cuesta, A. M., Álvarez-Vallina, L., & Fernández, L. A. (2015). Programming controlled adhesion of E. coli to target surfaces, cells, and tumors with synthetic adhesins. *ACS synthetic biology*, *4*(4), 463-473.

79. Vandenbroucke, K., De Haard, H., Beirnaert, E., Dreier, T., Lauwereys, M., Huyck, L., ... & Rottiers, P. (2010). Orally administered L. lactis secreting an anti-TNF Nanobody demonstrate efficacy in chronic colitis. *Mucosal immunology*, *3*(1), 49-56.

80. Pant, N., Marcotte, H., Hermans, P., Bezemer, S., Frenken, L., Johansen, K., & Hammarström, L. (2011). Lactobacilli producing bispecific llama-derived anti-rotavirus proteins in vivo for rotavirus-induced diarrhea. *Future microbiology*, *6*(5), 583-593.

81. Andersen, K. K., Strokappe, N. M., Hultberg, A., Truusalu, K., Smidt, I., Mikelsaar, R. H., ... & Marcotte, H. (2016). Neutralization of Clostridium difficile toxin B mediated by engineered lactobacilli that produce single-domain antibodies. *Infection and immunity*, *84*(2), 395-406.

82. del Rio, B., Redruello, B., Fernandez, M., Martin, M. C., Ladero, V., & Alvarez, M. A. (2019). Lactic acid bacteria as a live delivery system for the in situ production of nanobodies in the human gastrointestinal tract. *Frontiers in Microbiology*, *9*, 3179.

83. Powale, U. (2016). *Engineering Probiotic E. Coli With a Type III Secretion System for Targeted Delivery of Therapeutic VHH* (Doctoral dissertation).

84. Ruano-Gallego, D., Yara, D. A., Di Ianni, L., Frankel, G., Schüller, S., & Fernández, L. Á. (2019). A nanobody targeting the translocated intimin receptor inhibits the attachment of enterohemorrhagic E. coli to human colonic mucosa. *PLoS pathogens*, *15*(8), e1008031.

85. Virdi, V., Palaci, J., Laukens, B., Ryckaert, S., Cox, E., Vanderbeke, E., ... & Callewaert, N. (2019). Yeast-secreted, dried and food-admixed monomeric IgA prevents gastrointestinal infection in a piglet model. *Nature biotechnology*, *37*(5), 527-530.

86. Barta, M. L., Shearer, J. P., Arizmendi, O., Tremblay, J. M., Mehzabeen, N., Zheng, Q., ... & Picking, W. L. (2017). Single-domain antibodies pinpoint potential targets within Shigella invasion plasmid antigen D of the needle tip complex for inhibition of type III secretion. *Journal of Biological Chemistry*, *292*(40), 16677-16687.

Chapter 2

Plasmid vectors for *in vivo* selection-free use with the probiotic *E. coli* Nissle 1917

2.1. Abstract

Escherichia coli Nissle 1917 (EcN) is a probiotic bacterium, commonly employed to treat certain gastrointestinal disorders. It is fast emerging as an important target for the development of therapeutic engineered bacteria, benefitting from the wealth of knowledge of *E. coli* biology and ease of manipulation. Bacterial synthetic biology projects commonly utilize engineered plasmid vectors, which are simple to engineer and can reliably achieve high levels of protein expression. However, plasmids typically require antibiotics for maintenance, and the administration of an antibiotic is often incompatible with *in vivo* experimentation or treatment. EcN natively contains plasmids pMUT1 and pMUT2, which have no known function but are stable within the bacteria. Here, we describe the development of the pMUT plasmids into a robust platform for engineering EcN for *in vivo* experimentation, alongside a CRISPR-Cas9 system to remove the native plasmids. We systematically engineered both pMUT plasmids to contain selection markers, fluorescent markers, temperature sensitive expression, and curli secretion systems to export a customizable functional material into the extracellular space. We then demonstrate that the engineered plasmids were maintained in bacteria as the engineered bacteria pass through the mouse GI tract without selection, and that the secretion system remains functional, exporting functionalized curli proteins into the gut. Our plasmid system presents a platform for the rapid development of therapeutic EcN bacteria.

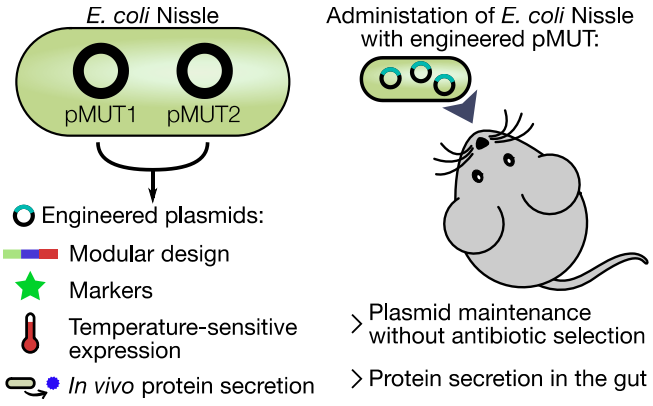


Figure 2.1: Graphical abstract.

2.2. Introduction

Escherichia coli Nissle 1917 (EcN) is a probiotic bacterium originally isolated from a particularly healthy soldier from World War I by the physician Alfred Nissle[1]. Since then, this bacterium has found significant use as a probiotic therapy, outcompeting pathogens in the gut[2] and thus protecting the host from infection. EcN has been at the forefront of probiotic genetic engineering[3], benefitting from the well-understood nature of *E. coli* biology, and from the many tools available to manipulate this organism. There are many projects working with engineered EcN[3], developing engineered therapeutic bacteria to tackle diseases in the gut like hyperammonemia[4], as well as outside the gut, such as for cancer detection and treatment[5,6].

In recent years, the gut microbiome has emerged as a critical factor for human health[7], however, the gut ecosystem remains a poorly understood system. One important approach to probe the gut microbiome is the development of engineered microbes that can sense and report on the conditions in the gut[8], as well as deliver therapeutic molecules into the gut environment[9]. Additionally, synthetic systems can provide insight into the behaviour of

engineered bacteria in the gut environment[10], aiding further engineering efforts. As such, genetic tools that simplify bacterial engineering both facilitate the study of gut health, and accelerate the development of sophisticated probiotic bacteria capable of sensing and treating gut disorders.

Synthetic biology projects typically utilize plasmid vectors, circular extrachromosomal DNA elements that can replicate within cells independently of the genome. Plasmids have many benefits: they are simple to manipulate, can be reliably transformed into *E. coli* cells, and can achieve high levels of gene expression due, in part, to a higher copy number than genomic DNA. Furthermore, several plasmids can be used in concert, allowing for modular assembly of complex synthetic genetic systems, as well as the simple independent testing of each plasmid in the system. An integral part of developing synthetic genetic systems is the iteration of prototypes in a design-build-test cycle[11], where during each cycle variants are tested to inform successive design iterations. Rapid and reliable genetic circuit construction and implementation is key for developing synthetic genetic systems, and plasmids offer an essential tool for this process. However, plasmid vectors also present a serious experimental limitation by requiring an antibiotic for selection and plasmid maintenance. In the context of *in vivo* therapeutic use in the gut, administration of an antibiotic is often incompatible with treatment and severely limiting to experiments as it induces drastic changes in the host microbiome[12].

Synthetic plasmids have been employed to engineer bacteria for *in vivo* use, however without antibiotic selection, significant plasmid loss has been observed[13]. Several plasmid maintenance strategies have been developed[14], including toxin-antitoxin systems[15], microcin mediated post-segregational killing[16], and auxotrophy[17]. However, such methods would require significant effort to optimize, and may themselves be burdensome to any

engineered genetic system. Given the limitations of plasmids, EcN engineering projects that require stable transformants often insert DNA directly into the chromosome. However, genomic manipulations are typically limited by poor transformation efficiencies in EcN, and involve time-consuming and cumbersome protocols, impeding the iteration of genetic circuit designs. Furthermore, common genomic incorporation protocols such as Lambda Red based methods can be inefficient and have limitations on insert length[18,19], further slowing or outright preventing the development of large multi-component synthetic genetic systems. Additionally, genomic incorporation limits recombinant DNA copy number to genomic copy number, making the achievement of high gene expression rates more difficult. Given the importance of rapid prototyping for the development of synthetic genetic systems, new paradigms are required to host synthetic DNA to facilitate the engineering of probiotic organisms.

Bacteria isolated from clinical samples often contain plasmids, including small cryptic plasmids that are maintained at high copy number despite containing little genetic information and conferring no apparent phenotype[20]. Many of these plasmids have no known function, although one study linked the presence of such small cryptic plasmids to phage resistance[21]. EcN contains two such cryptic plasmids, pMUT1 and pMUT2, which are stable within the bacteria, survive passage through the gut, and are used as targets to detect EcN in clinical PCR assays[22]. The pMUT plasmids do not confer any detectable phenotype, are not essential to EcN and do little to affect growth[23]. Furthermore, the pMUT plasmids do not present a metabolic burden to EcN, at least under laboratory conditions[24]. Whilst several projects have used pMUT plasmids to carry synthetic circuits[3], no systematic engineering attempt has been made to domesticate and characterize the efficacy of engineered pMUT plasmids *in vivo*.

In this work, we describe the systematic engineering of the *E. coli* Nissle 1917 cryptic plasmids pMUT1 and pMUT2 to create a series of plasmid vectors for use in the gut. We tested several sites on each plasmid to insert recombinant DNA cassettes containing selection and fluorescent markers, and characterized the gene expression in each case. We found that the native plasmids were not lost through transformation of an engineered variant, so we also developed a technique to remove the native plasmids through a CRISPR-Cas9 mechanism. We then added further functionality to these plasmid vectors: adapting and expanding a temperature sensitive expression system, as well as curli-based protein secretion to export proteins into the extracellular space. We then tested their use *in vivo*, finding that EcN retained the engineered pMUT plasmids during passage through the mouse GI tract, and that the plasmids were capable of secreting recombinant protein into the extracellular space of the gut.

2.3. Results

2.3.1. Revision of the pMUT cryptic plasmid sequences

EcN's cryptic plasmids were first documented by Hacker *et al.* in 2002, who published the sequences for pMUT1 and pMUT2 on the National Center for Biotechnology Information (NCBI) database with accession numbers A84793 and A95448 respectively[25]. Since then, 3 whole genome-sequencing projects for EcN have been uploaded to NCBI, with 2 fully assembled genomes. The first assembly, ASM71459v1 (Reister *et al.*[26]), resulted in a single sequence containing the chromosome and both plasmids, with the plasmids erroneously inserted multiple times within the chromosomal sequence. A later assembly, ASM354697v1, has a genomic sequence separate from the 2 cryptic plasmid sequences (labelled pNissle1 and pMUT2). Here,

the pNissle1 sequence contains both the sequence for pMUT1 and pMUT2 and is likely also an erroneous assembly. Also, the pMUT sequences from the whole genome sequencing projects differed from those originally uploaded, A84793 and A95448, which were sequenced using Sanger sequencing. As such, we could not find a correct pMUT1 sequence on NCBI and thus uploaded one for reference, NCBI accession number MW240712, and we refer to NCBI accession CP023342 for the correct pMUT2 sequence. We confirmed the pMUT sequences by Sanger sequencing the backbones of the pMUT-derived engineered vectors, finding the sequence traces aligned exactly with those derived from the whole genome sequencing efforts.

Figure 2.2a shows the plasmid maps and lengths of pMUT1 and pMUT2, and Supplementary Figure B1 shows the annotations as determined by the NCBI automated prokaryotic annotation pipeline [27] in greater detail. pMUT1 has a typical ColE1 origin of replication[22]. pMUT2 is 96.9% homologous to the pUB6060 plasmid from *Plesiomonas shigelloides*, which has been described as having a ColE2-like replication and ColE1-like mobilization loci[28]. Both plasmids contain *mob* genes involved in plasmid transfer, however both plasmids lack the full gene complement necessary for conjugation, and have previously been described as non-transferable[29]. Most of the putative proteins found on the plasmids have no known functions (Supplementary Figure B1), except for the *relB-relE* toxin-antitoxin system on pMUT2[30]. Toxin-antitoxin systems, often found on plasmids, are known to promote plasmid maintenance[15], so it is likely that these genes contribute to pMUT2 stability in EcN.

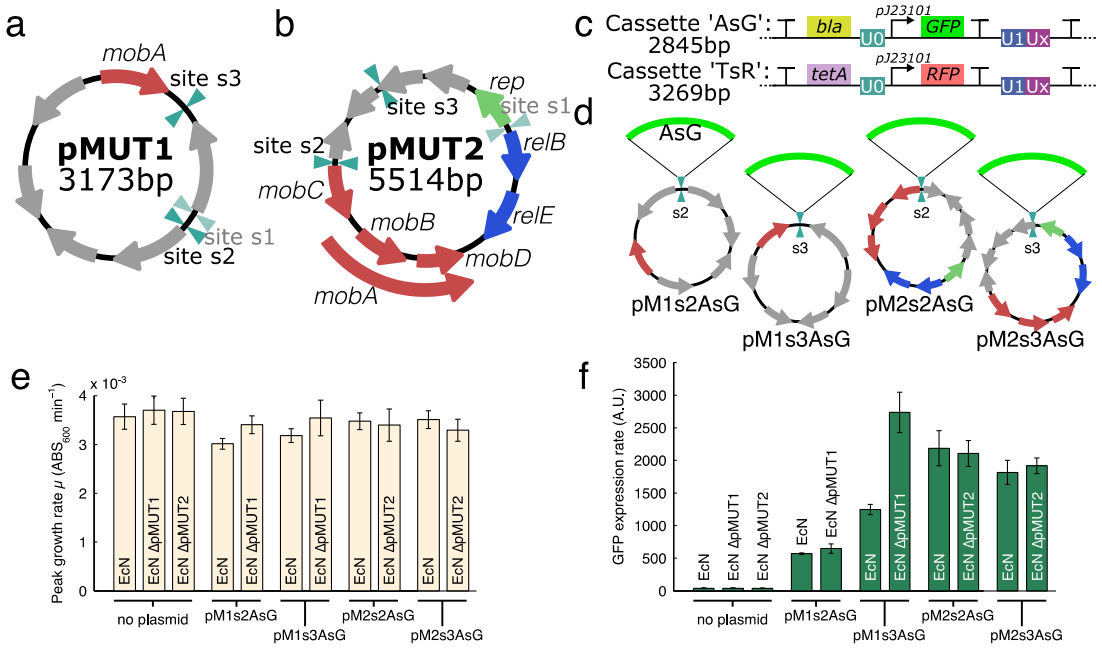


Figure 2.2: Plasmid maps of the native cryptic pMUT1 (a) and pMUT2 (b) plasmids in EcN with known genes labelled; also shown are the sites where we inserted recombinant cassettes, with unsuccessful sites greyed out. (c) Insulated characterization cassettes inserted onto the pMUT plasmids to produce engineered pMUT vectors, with cassette 'AsG' containing an ampicillin resistance gene and constitutively expressed GFP, and 'TsR' containing tetracycline resistance and constitutive RFP (characterized in Supplementary Figure B2). (d) For site characterization, we inserted the 'AsG' cassette onto the 2 sites on each pMUT plasmid, and characterized both the (e) bacterial growth rate and f) GFP expression rates. In each case, we characterized the performance of the engineered plasmids in both an unmodified EcN strain, and an EcN strain where the relevant native pMUT plasmid had been removed. Assays performed by Anton Kan and Siva Emani.

2.3.2. Engineering the pMUT plasmids

We began pMUT engineering by selecting 3 sites, s1 – s3, (Figure 2.2a-b) on each plasmid to insert an insulated cassette encoding antibiotic resistance and a fluorescent protein. The selected sites did not encode any known proteins or open reading frames and were away from the origin of replication in order to avoid disrupting any native functions. We kept the entire cryptic plasmid sequences as vector backbones, in order to maintain any features that may contribute to the stability of the plasmids.

We amplified the pMUT plasmids by PCR with primers (Supplementary Table B1) to act as the vector backbone onto which the cassette was inserted. For both pMUT1 and pMUT2, we attempted cassette insertion on 3 sites on the plasmid, but in both cases we could not assemble a plasmid for site 1 (s1) as we could not amplify the backbone by PCR despite trying two primer pair variants.

We tested the 2 successful insertion cassettes for gene expression (Figure 2.3c): 'AsG' which contained an ampicillin resistance gene (*bla*) and constitutively expressed superfolder GFP; and 'TsR', which contained a tetracycline resistance gene (*tetA*) and a constitutively expressed mCherry RFP. Both of these transcriptional elements were flanked by terminators to insulate the insertion cassette from transcription on the plasmid backbone and vice-versa. Furthermore, we added Universal Nucleotide Sequences (UNS) from Torella *et al.* [31] to the cassettes in order to allow for the rapid assembly of modular genetic elements. The UNS sequences are 40bp genetic segments that act as spacer elements without significant DNA structure or function. UNS sequences flanked each functional module we created, and are labelled as 'U#' in genetic circuit diagrams. The UNS sequence information can be found in Supplementary Table B2.

We cloned the inserts 'AsG' and 'TsR' into the sites on either pMUT1 or pMUT2 to obtain plasmids pMXsYAsG and pMXsYTsR, where X is either 1 or 2, referring to pMUT1 or pMUT2, and where Y is the insertion site number (Figure 2.2d). We tested the gene expression from each insertion site by measuring the cell density and GFP fluorescence in a kinetic run for each insertion site with plasmids pMXsYAsG. We performed these assays with the engineered pMUT plasmid transformed into either an unmodified EcN, or EcN where the corresponding native pMUT had been knocked out.

As reported before[23], EcN pMUT plasmid knockouts did not grow differently under lab conditions (Figure 2.2e). Growth rates were broadly similar in all cases with the engineered plasmids (Figure 2.2e), although in the pMUT1 derived vectors, the presence of the native pMUT1 reduced growth slightly ($p<0.001$ in both cases). We found that the insertion sites influenced GFP expression levels, with site 3 on pMUT1 and site 2 on pMUT2 giving the highest GFP expression levels (Figure 2.2f). The expression rates from pM1s3AsG further showed that the presence of the native plasmid altered gene expression from the engineered plasmid significantly. Since we would like our engineered vectors to be capable of high levels of gene expression, the high performing sites pM1s3 and pM2s2 were selected for further use. We also characterized RFP expression from pM1s3TsR and pM2s2TsR, finding a similar ratio of gene expression strengths to the GFP data (Supplementary Figure B2), suggesting that their relative rate of gene expression is independent of the protein expressed.

We found that transforming with the engineered plasmids did not displace the native plasmids. We tested for pMUT plasmids with DNA primers muta5, muta6, muta7, and muta8 (Supplementary Table B1), developed by Blum-Oehler *et al.*[22] to detect pMUT1 and pMUT2 in clinical samples. In a multiplexed PCR with these 4 muta primers, a 361 bp product is formed when pMUT1 is present, and a 429 bp product when pMUT2 is present (Figure 2.3a-b). Furthermore, we designed primers around the insertion sites on pMUT1 and pMUT2 to distinguish the native and engineered pMUT plasmids. We expected to find colonies in which the native pMUTs were knocked out through plasmid incompatibility – a process whereby two plasmids cannot stably coexist in the same bacterial cell line over multiple generations, typically occurring in plasmids containing similar or identical replication mechanisms. However, when unmodified EcN was transformed with an

ampicillin resistant engineered pMUT plasmid (pM1s3AsG or pM2s2AsG), and grown on selective media, colony PCR with primers around the insertion sites (primers pMXsY_chk_F and R) produced a short 200 bp product, indicating the presence of native plasmid (Figure 2.3e-h). This was despite the native and engineered plasmids having identical origins, only the engineered plasmids conferred antibiotic resistance. Since our data indicated that unmodified pMUT plasmid can impact the performance of the engineered plasmids, we developed a technique to rapidly remove the native plasmids prior to transformation with engineered pMUTs.

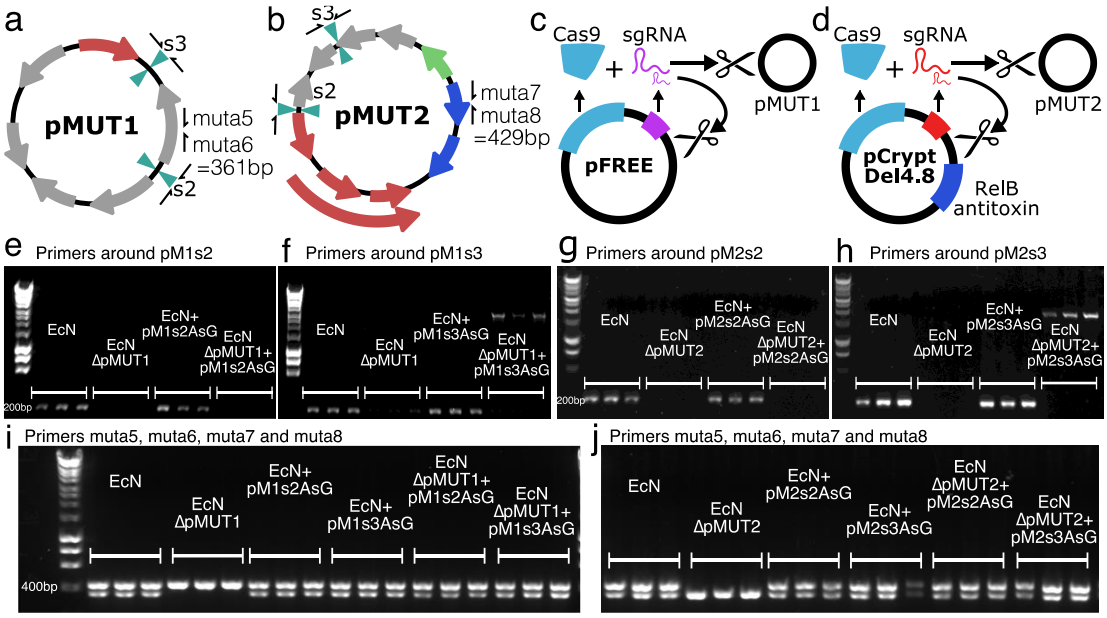


Figure 2.3: Curing native pMUT plasmids. EcN pMUT plasmids were assessed with primers around the insertion sites and primers muta5 and muta6 on pMUT1 (a), and muta7 and muta8 on pMUT2 (b). c) Plasmid pFREE cleaves pMUT1 through expression of Cas9 and gRNA targeting the colE1 origin of replication on pMUT1 and on pFREE itself. d) Similarly, pCryptDel4.8 targets the origin of pMUT2 and itself, and also contains a RelB antitoxin to disrupt the RelE-RelB toxin-antitoxin system on pMUT2. e-f) Agarose gels showing the results of colony PCRs around the insertion sites of the 'AsG' cassette, revealing that transformation with an engineered plasmid does not displace the native plasmid. i-j) pFREE and pCryptDel4.8 can cure EcN of native plasmids, and these can be replaced with engineered versions. In all agarose gel images, each condition is shown in triplicate, each lane representing a PCR result using a distinct bacterial colony. Assays performed by Anton Kan, Siva Emani, and Ilia Gelfat.

2.3.3. Curing the native pMUT plasmids

Since transforming EcN with engineered pMUTs did not displace the native plasmids, we required a strategy to remove them. Whilst pMUT curing strategies exist, they rely on plasmid incompatibility to knock out native plasmids[23], which our data indicates is not an immediate process and requires multiple weeks of streaking onto selective media, based on our experience. We therefore used a CRISPR-Cas9 strategy to cleave both native pMUT plasmids, based on the pFREE system of Lauritsen *et al.*[32]. The pFREE plasmid (Supplementary Figure B3a) contains a Cas9 protein and 4 guide RNAs (gRNAs) that target two sites on two common origins of replication, ColE1 and pSC101, in order to rapidly cleave and cure plasmids with those origins. The pFREE plasmid also harbours a variant of the ColE1 origin, and therefore also cures itself during this process.

We found that the pFREE plasmid cured pMUT1 (Figure 2.3c) with a success rate of around 60% (43 out of 73 colonies tested), but unsurprisingly did not cure pMUT2 as it does not have a ColE1 origin. To cure pMUT2, we redesigned guide RNAs to target locations on pMUT2 (Supplementary Table B3) to make the pCryptDel range of plasmids. We tested 3 gRNA pairs designed to cure pMUT2, however, all designs failed until we included the antitoxin gene *relE* found on pMUT2 onto the plasmids (Figure 2.3d). At each iteration of the design, we tested many pCryptDel variants for their ability to remove pMUT2, and identified one, labelled pCryptDel4.8, that cured EcN of pMUT2. Upon sequencing, we found that this plasmid targeted two sites on pMUT2 for Cas9-mediated cleavage, however it also contained an insertion mutation that altered one of the gRNAs targeting the ColE1 origin (Supplementary Figure B3). Variants without this mutation could cure pMUT1, but did not cure pMUT2 in any of the colonies tested. pCryptDel4.8 self-cured in all colonies tested, and cured pMUT2 with an efficiency of around 21% (24 out of 117 colonies tested),

however was very poor at curing pMUT1 (3 out of 97 colonies tested). In all cases of pMUT plasmid curing we tested, the pFREE and pCryptDel4.8 plasmids self-cured during the process. pFREE and pCryptDel4.8 can therefore be used to rapidly produce pMUT plasmid knockout strains through a single overnight growth step. Furthermore, the pFREE and pCryptDel plasmids can be used consecutively to create fully pMUT plasmid free EcN (Supplementary Figure B4).

2.3.4. Incorporating genetic modules for *in vivo* use: temperature sensing

The goal of our work was to make plasmid vectors for bacterial protein secretion in the gut. As such, there are several experimental challenges to both controlling and assessing synthetic genetic systems within the bacteria. Since the bacteria are in the gut of the host organism, they cannot be readily interrogated, and due to the complex environment of the gut, it is unlikely that bacteria behave as they do under laboratory conditions. This sets a severe limitation on genetic induction systems that require exogenous chemical inducers in the gut due to the difficulty of supplying a steady inducer concentration. Inducers are normally provided in a concentrated form in the water for the animal, so the effective concentration in the gut is not clear.

However, inducible systems are desirable to simplify cloning and *in vitro* propagation of DNA and bacterial strains, especially for genes that encode products that are toxic or stress-inducing to the bacteria. Synthetic genetic circuits typically require the bacteria to express heterologous proteins, and these can impose significant metabolic burdens on their host[33]. For constitutive high-expressing constructs, given a non-zero mutation rate, any defective mutants that relieve the metabolic burden will quickly come to dominate cultures due to faster growth. Therefore, for *in vitro* cloning and propagation it is desirable to use inducible systems to create an 'off' state where the synthetic system does not significantly

reduce fitness during culture propagation. Additionally, the uninduced state provides a further internal control in experiments that can provide valuable insight into the performance of the genetic system.

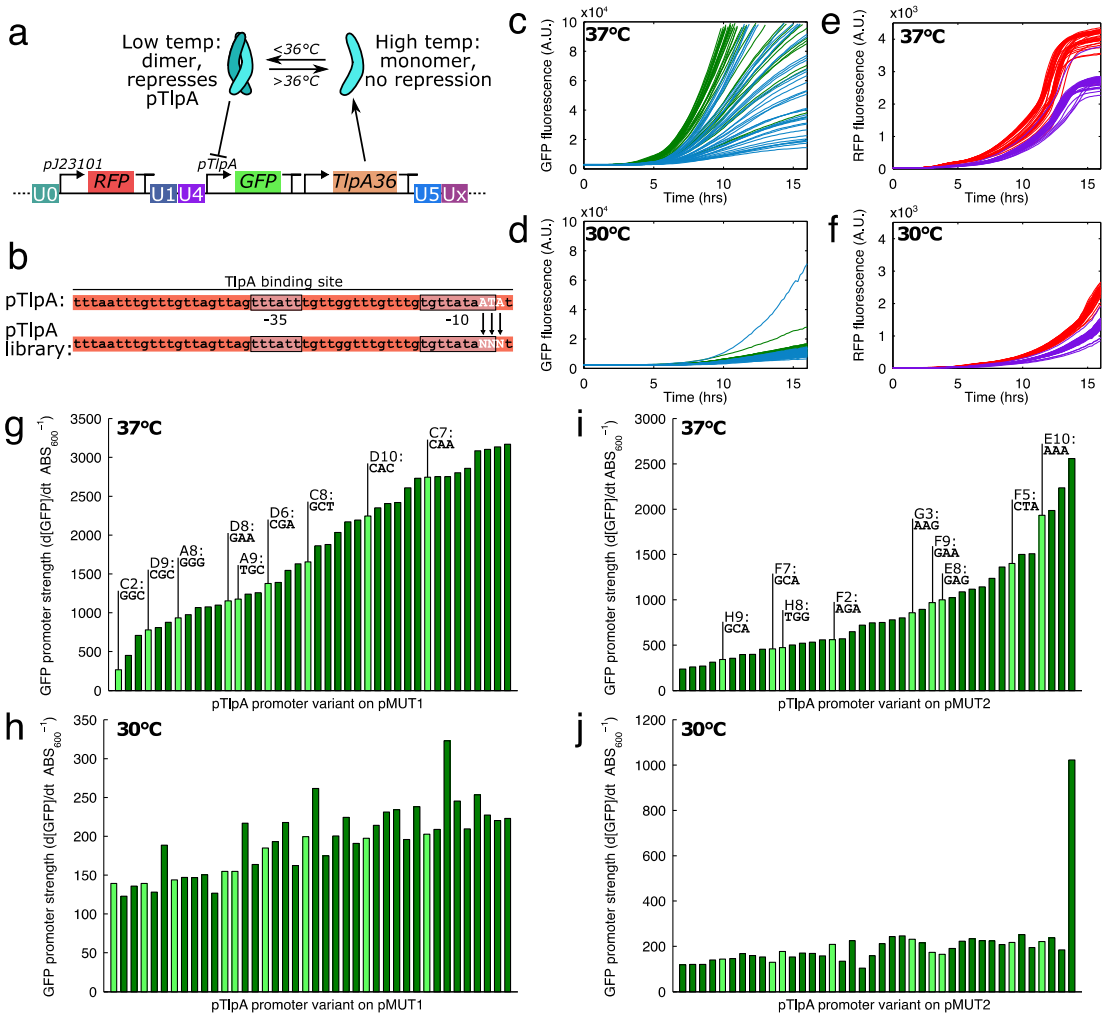


Figure 2.4: (a) Temperature sensitive expression was achieved with the TlpA36 protein, which dimerizes, binds, and represses the pTlpA promoter at temperatures below 36°C. (b) To make a library with various promoter strengths, the pTlpA promoter was modified to contain 3 variable nucleotides near the -10 region of the promoter. Variable GFP expression strengths at 37°C (c) and 30°C (d) from the pTlpA library, with green curves showing 40 pM1s3AsR_TS* variants, and blue showing 40 pM2s2AsR_TS* variants. (e-f) In contrast, RFP was expressed by a constitutive J23101 promoter and RFP expression rates were not as variable as for GFP. When promoter strengths were quantified at 37°C for the pMUT1 (g) and pMUT2 (i) engineered vectors, we found a range of strengths, and 9 promoters throughout the range were chosen for sequencing and further development (highlighted in

lighter green). (h-j) In general, gene expression from the pTlpA* promoters was reduced at 30°C by at least tenfold. Assays performed by Anton Kan and Siva Emani.

We implemented a temperature-sensitive gene expression system from Piraner *et al.*[34], based on the promoter pTlpA and repressor protein TlpA36. TlpA36 forms a dimer at temperatures below 36°C, and this dimer binds pTlpA and prevents gene expression (Figure 2.4a). At temperatures above 36°C, the repressor dimer is unstable and does not prevent gene expression. As such, this is an ideal system to provide constitutive high gene expression *in vivo*, as both human and mouse body temperatures are around 37°C, whilst *in vitro* the bacteria can be grown at 30°C. To characterize the temperature dependent gene expression, we designed a synthetic ratiometric construct containing pTlpA driving sfGFP, and the constitutive pJ23101 (BioBricks registry) driving mCherry, an RFP (Figure 2.4a).

Above a certain critical temperature, the pTlpA promoter is active and acts constitutively. We mutated the promoter to generate a library with varied expression strengths. The promoter variant could then be selected for a transcriptional unit of interest in order to optimize gene expression. TlpA, from which TlpA36 was derived, binds to the entire pTlpA promoter[35] (Figure 2.4b), so to minimally disrupt the repressor-promoter interaction we limited our mutations to the edge of the promoter region. We found the -35 and -10 regions of the promoter with the BPROM software[36], and designed the 3 mutations in the -10 region of the pTlpA promoter, a region important to RNA polymerase binding and transcription strength. We therefore designed DNA oligos containing 3 freely varying nucleotides, and used them to assemble the 'sR_TS' (i.e., "RFP, temperature sensitive") circuit shown in Figure 2.4a on both the pM1s3A and pM2s2A backbones. The assembled plasmids were then transformed into an *E. coli* cloning strain, Mach1, for an initial screening and characterization. Transformants were initially screened for GFP on LB agar plates grown

at 37°C to select variants with a range of expression strengths, and 40 constructs were chosen each for pM1s3A and pM2s2A. Constructs were labelled by the microwell within the plate, with pM1s3AsR_TS-A1 to D10 for the engineered pMUT1 constructs and pM2s2AsR_TS-E1 to H10 for the pMUT2 constructs. The cells bearing the selected constructs were then monitored during growth in the plate reader, where we quantified the GFP, RFP fluorescence and absorbance at 600 nm at both 37°C and 30°C to identify their strength and inducibility (Figure 2.4c-f). We found that the GFP expression strengths varied significantly between constructs at 37°C, whereas RFP expression remained similar in each case and did not vary as much with temperature.

We found the promoter strengths of the pTlpA variants (labelled pTlpA-A1 to pTlpA-H10) by calculating the amount of GFP produced per unit time and per cell (Figure 2.4g-j). We found that the promoters were weaker by around an order of magnitude at 30°C compared to 37°C, and when induced, covered a wide range of expression strengths on both pMUT vectors. We selected 9 of these constructs to cover the range of expression strengths each on the pM1s3A and pM2s2A backbones for sequencing and further use. The selected constructs are highlighted in light green on the Figure 2.4g-j, and the mutant nucleotides are shown as labels above. A table with the promoter sequence and their associated expression strengths can be found in Supplementary Table B4.

We then characterized the performance of the pM1s3AsR_TS* and pM2s2AsR_TS* constructs in EcN, in each case measuring the engineered pMUT construct performance in the absence of the native cryptic plasmid. We found that we could not transform some of the pM2s2AsR_TS* constructs into EcN ΔpMUT2 cells, and thus we could only characterize 4 of the pMUT2 derived constructs. The characterization data from the *E. coli* Mach1 cloning strain was broadly indicative of performance in EcN (Figure 2.5a), although the pM2s2A

constructs in particular did not fully match their behaviour in Mach1. The pTlpA* promoters covered a range of expression strengths when induced, and had significantly less expression at 30°C (Figure 2.5b). By contrast, constitutive RFP expression from each construct was similar, and did not vary as much with temperature (Figure 2.5c).

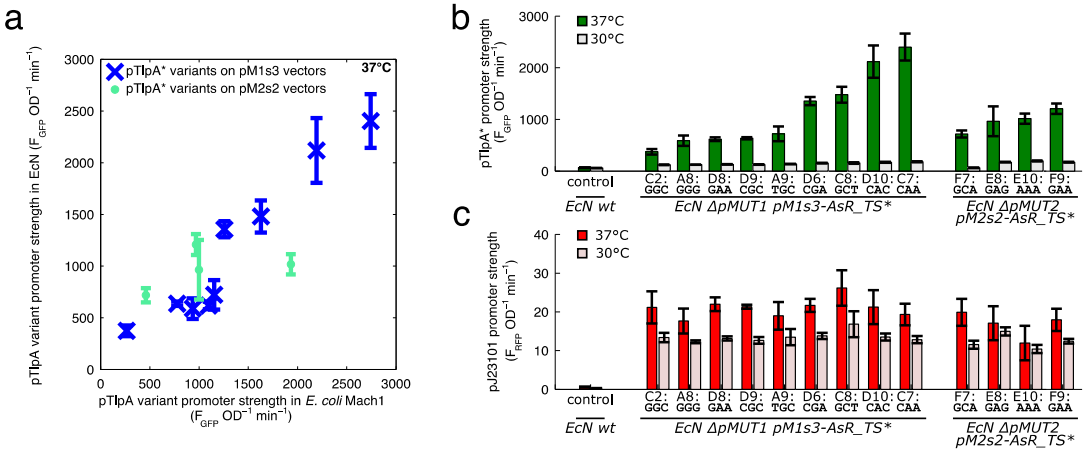


Figure 2.5: (a) Comparison of the temperature sensitive gene expression circuit in EcN and in *E. coli* Mach1 strains. (b) GFP expression from the pTlpA library in EcN, showing a range of expression strengths. (c) Constitutive RFP from the 'AsR_TS*' cassettes from the engineered pMUT plasmids in EcN. Assays performed by Anton Kan and Siva Emani.

2.3.5. Curli secretion from engineered pMUT vectors

Many proteins and peptides have therapeutic potential in the gut[37], and as such the secretion of such peptides into the extracellular space from EcN inhabiting the gut is an attractive approach to therapy. Curli are well-characterized bacterial extracellular matrix proteins, secreted natively by *E. coli* using dedicated machinery[38] to form robust fibers. Engineered curli systems are emerging as a versatile platform for custom protein materials, as they are capable of tolerating mutations and fusions to functional protein domains, and consequently they are being developed as gut therapeutics[39].

To express curli, we used a synthetic *csgBACEFG* operon[40], which encodes the major and minor curli fiber subunits, *csgA* and *csgB*, and the secretion machinery necessary for transport from the periplasm to the extracellular space in *csgEFG*. CsgC prevents intracellular CsgA polymerization, which would be toxic to the bacterium[41]. We fused the CsgA monomer to an E-tag epitope tag in a 37 aa flexible linker (Figure 2.6a) to enable detection with anti-E-tag antibodies, and this cassette was 'csg-Etag'. In order to further demonstrate a functional curli variant, we also produced constructs where CsgA was fused to a GFP nanobody (NbGFP)[42]. Nanobodies, also known as VHH domains, are single chain antibody fragments[43], capable of binding tightly to a specific antigen. In this case, the *csgA:NbGFP* fusion should bind GFP, and due to the insolubility of the curli material, should remove purified GFP from solution. The *csgA:NbGFP* fusion in cassette 'csg-Etag-NbGFP' also encoded an E-tag in a 37 aa flexible linker between the *csgA* and *NbGFP* sequences, and the *NbGFP* was followed by a 6xHis tag (Figure 2.6a).

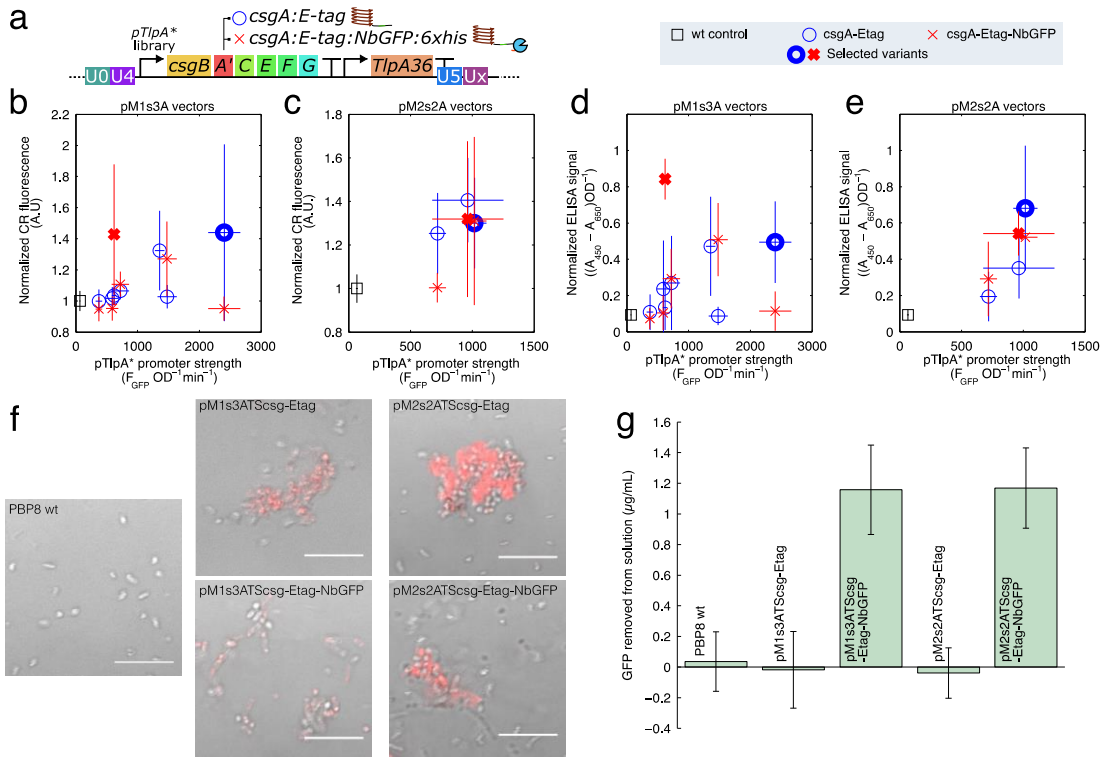


Figure 2.6: (a) Diagram of the temperature sensitive curli production construct, containing a pTlpA promoter variant driving the expression of a synthetic curli operon *csgBACEFG*, where the *csgA* sequence is appended with an E-tag epitope tag (labelled cassette 'csg-Etag'), or an E-tag and a GFP nanobody sequence with 6xHis (cassette 'csg-Etag-NbGFP'). The temperature sensitive pTlpA* promoter variants were all used to generate pM1s3ATScsg-# variants (b and d) and pM2s2ATScsg-# (c and e) variants. These variants were assayed with a Congo Red (CR) curli assay (b and c), where CR dye stains the curli proteins and becomes fluorescent, and an anti-Etag filtration ELISA (d and e). In panels b-e, thick markers represent the variants selected from the promoter library for subsequent use. (f) Representative confocal micrographs of bacterial cultures harbouring the selected plasmids with temperature inducible curli grown at 37°C in the presence of CR, with the red CR fluorescence overlaying a brightfield image. Scale is 10 μm. (g) Curli fused to GFP nanobodies (NbGFP) was able to remove a significant amount of sfGFP from a 4 μg/mL solution of purified sfGFP in PBS. Assays performed by Anton Kan and Siva Emani.

Overexpression of the *csgBACEFG* operon can be toxic to cells, and as such the expression strength requires significant tuning to obtain a high yield of curli fibers. We therefore used the pTlpA promoter library to express a synthetic curli operon to identify an optimal promoter strength. In total, we were able to generate 8 of each 'csg-Etag' and 'csg-Etag-NbGFP' on pM1s3A vectors, and 3 each on pM2s2A vectors. For each pTlpA*-curli construct, we characterized the curli production using the Congo Red (CR) fluorescence method[44] (Figure 2.6b-c), as well as by a filtration ELISA method with anti-E-tag antibodies (Figure 2.6d-e). Both methods showed similar results, with certain combinations of promoter and *csgA* variant producing significant yields of curli material. For the *csgA:Etag* constructs, higher promoter strength generated the higher curli yields, and we chose the highest expressing constructs to make plasmids pM1s3ATScsg-Etag and pM2s2ATScsg-Etag, which used the pTlpA-C7 and pTlpA-E8 promoters respectively. By contrast, the 'csg-Etag-NbGFP' constructs had peak expression at intermediate promoter strengths, with the chosen plasmids pM1s3ATScsg-NbGFP and pM2s2ATScsg-NbGFP containing the pTlpA-D8 and pTlpA-E10 promoters. In all cases, the temperature sensitive curli constructs expressed GFP poorly at 30°C (Supplementary Table B4).

At 37°C, the final temperature sensitive curli expression constructs produced curli, which caused the bacterial cultures to form aggregates that were fluorescent upon the addition of CR (Figure 2.6f). Additionally, at 37°C, bacteria with the *csgA:NbGFP* fusions successfully bound and removed purified GFP from a solution of purified GFP, demonstrating the desired function of the nanobody (Figure 2.6g).

2.3.6. Performance of engineered pMUTs in the mouse gut

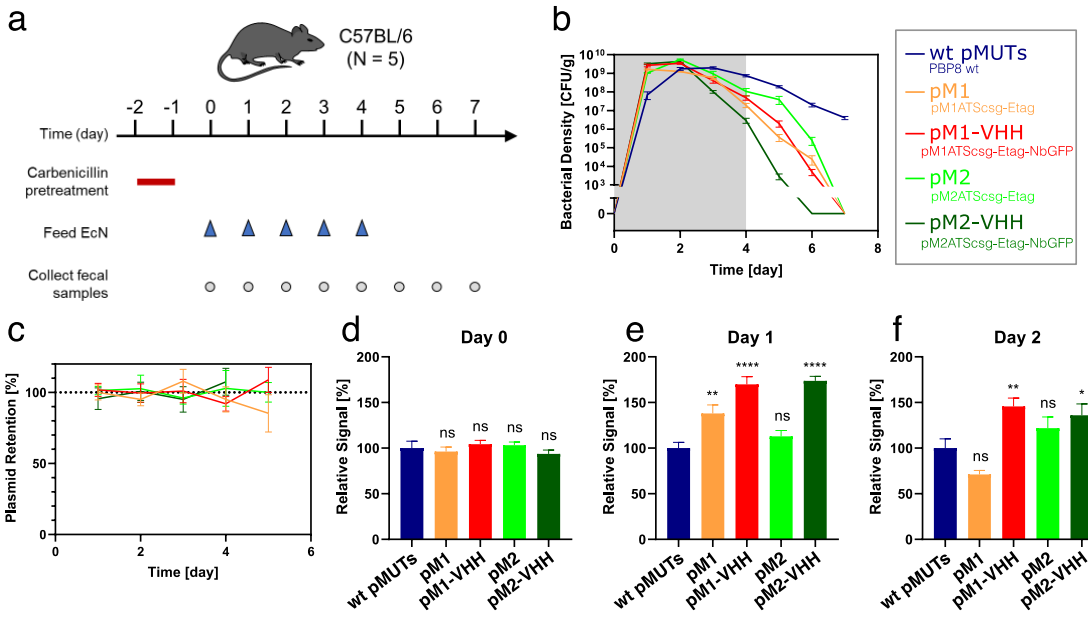


Figure 2.7: Engineered pMUTs in the mouse gut. (a) Timeline of *in vivo* study. (b) Bacterial density of PBP8 over time, as measured by CFU counts from fecal samples plated on LB agar with Cm. (c) Plasmid retention over time. (d-f) Relative *in vivo* protein expression levels from fecal filtration ELISA on days 0, 1 and 2 (d, e and f, respectively). At each day, engineered pMUT conditions were tested against a WT pMUT control by one-way ANOVA, followed by pairwise Welch's t-test. ns – not significant, * $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$. All data are represented as mean ± SEM. Assays performed by Ilia Gelfat and Anton Kan.

One of the original motivations for this work was to address and improve retention rates of synthetic plasmids in bacteria within the mouse gut. In a preliminary experiment testing engineered EcN in the mouse gut, we found that EcN harboring engineered synthetic plasmids were lost during passage through the gut without selection. In this experiment,

we fed mice PBP8 cells (EcN Δ*csgBACEFG*::cat(CamR)) transformed with either plasmid pKAG[45], a pSB4K5 based plasmid containing constitutively expressed sfGFP, or pL6FO[44], a similar synthetic plasmid with an IPTG inducible *csgBACEFG* operon (Supplementary Figure B5a). The engineered bacteria were administered to the mice on day 0 of the experiment (Supplementary Figure B5b), and we tracked both the overall PBP8 population and the plasmid bearing population in fecal samples over the subsequent days. As we were selecting for the PBP8 bacteria by treating with chloramphenicol, we found that the PBP8 persisted in the gut in all cases after administration (Supplementary Figure B5c). However, we found significant plasmid loss for all synthetic plasmids (Supplementary Figure B5d), which we tested by challenging with kanamycin. On day 1 after administration, pL6FO was only present in only around 15% of the population when the curli operon was in the IPTG induced state. Furthermore, for plasmid pKAG, and pL6FO with IPTG induced curli operon, the plasmid-bearing bacteria were not found in the gut after 5 days.

We sought to test the plasmid retention of our engineered pMUTs after passing through the mouse gut. Additionally, we were interested in determining the ability of our plasmid system to produce and secrete proteins in an *in vivo* context, as this feature is key to therapeutic peptide delivery in the gut. Bacterial gene expression in a mammalian gut significantly differs from expression under laboratory conditions[46], and as such *in vitro* characterization is unlikely to be representative of *in vivo* functionality.

Typically, it is difficult to assess the gene expression of engineered bacteria in the gut, because the bacteria must usually be grown *in vitro* after isolation from fecal samples, which disrupts any measurement of *in vivo* gene expression. Direct detection of heterologously-produced proteins in fecal samples is similarly challenging. For most proteins and affinity tags, proteolytic degradation by intestinal proteases is likely to

significantly reduce any measurable signal. This is particularly problematic considering the high background signal one can expect from a complex biological medium such as feces. These experimental limitations were, in large part, what motivated us to test the pMUT system using curli fibers and VHH domains. In addition to the potential utility of these proteins, both curli and nanobodies are known for their resistance to the harsh conditions[47,48], thereby increasing the likelihood of their detection in fecal pellets.

We designed an experiment to test the retention of the engineered pMUTs *in vivo*, as well as the expression of protein through the plasmid system within the mouse gut. Four plasmids were tested, expressing either cassette 'csg-Etag' or 'csg-Etag-NbGFP' on pM1s3ATS* or pM2s2TAS* vectors. In each case, PBP8 cells (EcN Δ*csgBACEFG*::cat(CamR)) were used, with the native pMUT knocked out whenever the corresponding engineered version was present. As a negative control, we used PBP8 harboring both wild-type pMUTs with no engineered plasmids, making for a total of 5 experimental cohorts. We labelled the conditions: 'wt pMUTs' for the control; 'pM1' for PBP8 ΔpMUT1 pM1s3ATScsg-Etag; 'pM1-VHH' for PBP8 ΔpMUT1 pM1s3ATScsg-Etag-NbGFP; 'pM2' for PBP8 ΔpMUT2 pM2s2ATScsg-Etag; 'pM2-VHH' for PBP8 ΔpMUT2 pM2s2ATScsg-Etag-NbGFP. Each of the five bacterial strains were administered to C57BL/6 mice (n = 5).

The mice were fed bacterial suspension daily for 5 days and monitored for 3 additional days after cessation of bacterial administration (Figure 2.7a). Each day, fecal pellets were collected for colony counting and protein detection. Like most human *E. coli* isolates, EcN is a poor colonizer of the mouse gut[49], though it can transiently persist in mice pre-treated with antibiotics. Therefore, carbenicillin was given two days prior to bacterial feeding, in order to allow the engineered EcN to reach high density, with the antibiotic administration

lasting 24 hours to avoid artificially selecting for the engineered plasmids. As a result, EcN density gradually dropped over the course of the experiment, likely due to the recovery of native mouse microbiome (Figure 2.7b). All the engineered EcN were cleared from the mice by day 7, unlike the wt control, and we hypothesized that this was due to the increased fitness of the unmodified bacteria. When characterized *in vitro* (SI Figure 2.7), the unmodified EcN grew significantly faster than any engineered pMUT variant, suggesting that the burden of the recombinant gene expression reduced fitness.

Each fecal pellet was plated on two types of selective plates: chloramphenicol (Cm), selecting for PBP8 irrespective of plasmid presence or identity; and chloramphenicol with carbenicillin (Cm+Carb), which selected specifically for PBP8 with an engineered plasmid. Plasmid retention rates were calculated as the ratio of Cm+Carb to Cm colony counts. All four engineered pMUT cohorts showed no plasmid loss during GI transit, with none of the retention rates differing significantly from 100% (Figure 2.7c).

Protein expression was tested via fecal filtration ELISA, modified from a previous protocol[39]. In both engineered pMUT1 and pMUT2, significant levels of E-tagged curli fibers were detected (Figure 2.7d). Since the mice were fed EcN grown at 30°C, there was no curli expression prior to feeding, so this result demonstrated the ability of the engineered pMUT system to express and secrete proteins *in vivo*. Interestingly, in both plasmids, the nanobody-containing constructs produced a higher signal than their nanobody-free counterparts. We suspect this may be due to the robust CsgA and NbGFP protein domains flanking and protecting the E-tag from proteolysis when in the gut lumen.

2.4. Discussion

In this work we developed plasmid vectors based on the *E. coli* Nissle 1917 pMUT cryptic plasmids, and characterized their performance in the mouse gut. Our work developed a simple method to remove the native pMUT plasmids, and generated reliable pMUT plasmid vectors capable of secreting a functional curli material within the mouse gut without plasmid loss. Furthermore, our pMUT-based plasmid vectors simplified *in vivo* experiments by forgoing the need for antibiotics for plasmid maintenance or inducers for gene expression through temperature sensitive circuits.

The pMUT plasmids have no known function, but are stable within EcN during passage through the gut, and can thus act as vectors for recombinant DNA. Whilst previous studies have used the pMUT plasmids[3], and shown their high plasmid retention *in vitro*[24], their *in vivo* efficacy had never been systematically characterized. Through our attempts to cure the native pMUT plasmids, our results suggest that pMUT2 stability in EcN is improved by the RelB/RelE toxin-antitoxin system, as we could not cure pMUT2 without expressing the antitoxin gene from our pCryptDel4.8 plasmid. This demonstrates how the approach of building and testing genetic tools not only creates useful systems, but also provides insight into the underlying biology.

Despite the common use of plasmids in the development of engineered microbes, they are not typically utilized in clinical applications, where exogenous genetic sequences are instead incorporated into the chromosome of the chassis organism. A major reason for this is plasmid loss, and this phenomenon severely limits the efficacy of plasmid based genetic systems *in vivo*. Antibiotics, commonly used for plasmid maintenance *in vitro*, are incompatible with many *in vivo* assays, as they disrupt the native microbiota as well as any

pathogens, and additionally raise significant questions regarding the effective antibiotic dose. Engineered plasmid maintenance strategies could address this issue, but such methods have so far not fully overcome plasmid loss[16], and would currently require significant development and optimization. The addition of engineered plasmid retention would also create a further metabolic burden from the plasmid, potentially reducing the efficacy of other synthetic genetic circuits. The pMUT plasmids, by contrast, are in some sense already optimized for EcN, as they are stable and do not impose any noticeable burden. Whilst it is possible that some recombinant inserts may hinder their stability, we did not observe any loss of the engineered pMUT plasmids, despite adding synthetic circuits that imposed a significant reduction in growth when induced.

A further concern for plasmid based engineering is horizontal gene transfer (HGT), whereby a plasmid with an antibiotic resistance gene or virulence factor would run the risk of being introduced into the host microbiome[50]. While such concerns are valid for most synthetic plasmids, the unique features of the engineered pMUT system could address these limitations. Most prominently, the absence of antibiotic selection could eliminate the possibility of spreading resistance genes, as the resistance gene can be excised from any engineered bacterium through a recombinase. Furthermore, the presence of these plasmids in wild-type EcN suggests that the risk posed by any sequence found natively on the plasmid is negligible. Indeed, the safety profile of EcN over decades of probiotic use implies that HGT of pMUT-encoded genes is either exceedingly rare, relatively harmless or both. Lastly, some have proposed utilizing HGT as a tool for *in situ* microbiome engineering[51]. As a selection-free, probiotic-derived plasmid system, the pMUT platform could prove a valuable addition to the toolbox of this emerging microbial intervention strategy. Thus, while the pMUT plasmids could indeed be utilized for the research and development of

engineered strains, they could also open the door to plasmid-based production of therapeutics *in vivo*, in both clinical and preclinical settings.

There are several benefits to using engineered pMUT plasmid vectors compared to genomic incorporation. The first is speed and reliability, since plasmid assembly and transformation are the only steps required for the production of an engineered EcN strain, and this can be done in several days. This can facilitate the rapid construction and development of probiotic bacteria, speeding the development and optimization of prototypes. A further benefit is the ability to incorporate relatively large recombinant genetic constructs with ease. Indeed, one of the largest constructs we made was around 13 Kbp (pM2s2ATScsg-NbGFP), incorporating over 7 Kbp of recombinant DNA. Furthermore, both engineered pMUT1 and pMUT2 plasmids could be used simultaneously to house synthetic DNA, allowing for the incorporation of even larger and more complex synthetic DNA systems.

A major benefit to simplifying the process of bacterial engineering is the ability to rapidly and reliably generate multiple variant strains, and thus screen and optimize genetic circuits of interest. The pTlpA promoter library in this case demonstrated how even a relatively small functional change, such as the addition of a fusion protein, can require the redesign of regulatory elements within genetic circuits for optimal function. Here, the addition of an anti-GFP nanobody required a weaker promoter for curli expression compared to unmodified CsgA-Etag, suggesting that the nanobody reduced secretion efficacy, likely through the toxicity of expression and secretion. However, the weaker expression did not reduce overall curli production in the nanobody constructs, suggesting that curli production was not limited by the expression of the other genes in the *csgBACEFG* operon.

In our *in vivo* experiments, we observed slower clearance of the WT pMUT control strain compared to those expressing proteins through engineered pMUTs (Figure 2.7b). We believe this reflects the added metabolic burden imposed on the cells by overexpression of heterologous protein, rather than any feature of the engineered plasmids, as our previous work yielded similar trends with different plasmid systems[39]. In addition, as was the case in the aforementioned study, bacterial density and protein expression varied between the different conditions, indicating these factors depend on the specific proteins being expressed. Praveschotinunt *et al.*[39] also demonstrated that PBP8 and WT EcN do not differ substantially in their *in vivo* behaviour, and that strains expressing wild-type curli fibers can exhibit similar bacterial densities to those producing GFP. These findings suggest that the performance observed by the strains in this work is unlikely to be specific to curli. Taken together, such observations support the compatibility of engineered pMUTs with *in vivo* expression of a variety of proteins, though the expression strength would have to be adjusted to achieve optimal results for each desired application – as would be the case for any other expression platform, be it genomic integration or plasmid-based.

2.5. Conclusion

In this work we have harnessed native cryptic plasmids to create a robust genetic platform for engineering probiotic *E. coli* Nissle 1917 bacteria. Whilst *E. coli* is not a major component of the human or mouse microbiome, it is often present at sites of inflammation[52]. As such, *E. coli* is capable of playing an important therapeutic role, both by competing with pathogenic bacteria as well as being able to deliver therapeutic compounds to where they are needed.

It is becoming increasingly clear that the state of the gut microbiome has important ramifications for human health, and there are many unanswered questions about the role of the microbes. In this work we have developed a reliable genetic platform to host synthetic DNA for *E. coli* Nissle. Our platform simplifies research, facilitating new experiments to investigate the gut microbiome, and speeds the development of therapeutic engineered bacteria that can be deployed clinically.

2.6. Materials and Methods

2.6.1. DNA cloning

All plasmid assembly was performed using Gibson Assembly, with the exception of the pCryptDel# plasmids, where the gRNA array was assembled using Golden Gate assembly due to many repeats in the DNA sequence. Custom DNA oligos were ordered from Integrated DNA Technologies (IDT) and used in PCR with Q5 polymerase (New England Biolabs) to create amplicons for subsequent Gibson assembly. DNA purification from PCR was done with ZymoClean Gel DNA Recovery Kit (ZymoGen). DNA assembly products were transformed into chemically competent *E. coli* Mach1 cells (Thermo Fisher Scientific) and plated onto LB Agar plates with appropriate antibiotics.

DNA libraries were generated by designing degenerate bases at selected locations on DNA oligos, flanked by 25 bp of the unmodified sequence. The resulting DNA was synthesized (IDT) used as primers to make amplicons for plasmid assembly. The resulting pool of assembled plasmid variants was transformed into Mach1 cells and plated onto 10 plates. After overnight incubation at 37°C, the plates were imaged for GFP and RFP fluorescence in a FluorChem M Imager (Protein Simple), and colonies were selected.

2.6.2. Colony PCR

Assessment of cryptic plasmids was done by colony PCR using 25 µL reactions with Quick-Load Taq PCR mix (New England Biolabs) following the manufacturer's instructions. After the PCR, the reactions were added to a 1% agarose TAE gel with SybrSafe DNA stain and ran in a gel electrophoresis setup (constant 120V, 35 mins). Gels were then imaged in FluorChem M Imager (Protein Simple).

2.6.3. Bacterial culture

E. coli bacteria were grown in LB Miller media during plasmid preparation and genetic circuit characterization. For characterization assays, starter cultures of the appropriate bacterial cultures were grown overnight in LB media in a shaking incubator. For all temperature sensitive constructs, started cultures were grown at 30°C, whereas we used 37°C for all other constructs. Unless explicitly indicated otherwise, all characterization was done at 37°C.

Kinetic plate reader assays were performed by diluting starter culture 1:1000 into the appropriate selective media. We then added 200 µL of the inoculated culture into the wells of black, clear-bottom, 96-well plates (655090, Greiner Bio-One). The plates were then grown in a Synergy HT plate reader (BioTek), reading absorbance (600 nm), GFP (excitation: 485/20 nm, emission: 528/20 nm), RFP (ex: 590/20, em: 645/40). Reads were taken every 10 minutes for 16 hours, and plates were shaken continuously outside of reading (Double Orbital, 548 cpm (2 mm)).

2.6.4. Plasmid curing

In order to cure Nissle and any derived strains of cryptic plasmids, they were transformed with plasmids pFREE or pCryptDel4.8, in order to cure pMUT1 or pMUT2 respectively.

After transformation, cells were grown overnight in liquid LB media with 50 μg/mL kanamycin. Then, the overnight culture was diluted 1:1000 into fresh LB media supplemented with 50 μg/mL kanamycin, 0.2% rhamnose and 0.43 μM anhydrous tetracycline (ATC), and grown overnight at 37°C. After 24 hours, the culture was streaked out onto several LB agar plates without antibiotics and these were left to grow overnight. Then, the colonies were assessed by colony PCR with primers muta5, muta6, muta7 and muta8 to find colonies that had been cured of cryptic plasmids.

2.6.5. Growth and gene expression characterization

Data from kinetic plate reader runs was initially cleaned by subtracting the background signal and smoothing the time courses for all fluorescence and absorbance data. Growth rates were found by fitting the absorbance curves to a Gompertz model, and subsequently extracting the peak growth rate. Promoter strength was quantified from kinetic fluorescence data by first finding the gradient of the fluorescence signal, normalizing this to the absorbance signal, resulting in a per cell measure of fluorescent protein production per unit time. Promoter strength was then quoted to be the average of this term for an hour around peak exponential phase.

2.6.6. Curli measurement

Curli was measured by the CR method outlined in Kan *et al.*[44]. Bacterial starter cultures were grown overnight in LB and the relevant antibiotics at 30°C. Then, we inoculated selective LB media 1:1000 with starter culture, and placed 300 μL into 1 mL deep well plates (780210, Greiner Bio-One) in a shaking incubator (900 rpm (1 mm)) at either 30°C or 37°C. 0.025% CR was added to the media upon inoculation. After 24 hours of growth, 200 μL of each well was transferred into black, clear bottom plates and the absorbance (600 nm)

and CR fluorescence (ex: 525 nm, em: 625 nm) was read in a Synergy HT plate reader. The results were then normalized to the host strain without engineered plasmids.

Curli production was also measured by whole cell filtration ELISA to measure the E-tagged CsgA proteins. 80 µL bacterial overnight cultures were added into each well of a 96-well filter plate in triplicate (0.22 µm pore size, Multiscreen-GV, Merck/Millipore Sigma). Samples were vacuum-filtered, and washed in TBST (TBS, 0.1% Tween-20), and blocked for 1.5 hours at 37°C with 1% bovine serum albumin (BSA) and 0.01% H_2O_2 in TBST. After additional washing, samples were incubated with HRP-conjugated goat polyclonal E-tag epitope antibody (Novus Biologicals) for 1.5 hours at room temperature (1:5000 in TBST). Following additional washes in TBST, 100 µL Ultra-TMB ELISA reagent (Thermo Scientific) was added to each well and covered with aluminium foil to protect from light. After approximately 15-25 minutes of incubation at room temperature, the reaction was stopped using 50 µL per well of 2 M sulfuric acid. 100 µL were transferred from each well into a 96-well plate and absorbance was measured at 450 nm and 650 nm. The signal was calculated by subtracting the 650 nm absorbance value from the 450 nm absorbance value.

2.6.7. GFP sequestration assay

To test the function of the GFP nanobody, bacterial cultures were first grown overnight at stated temperatures in LB media with appropriate antibiotics. They were then pelleted at 3000g for 10 minutes, washed once in PBS, and resuspended in a solution of PBS containing 4 µg/mL purified sfGFP. The solutions were then left in a rotating mixer for an hour at room temperature, then centrifuged again at 3000g for 10 minutes. The supernatant GFP signal was then measured in the plate reader, and compared to the fluorescence of the initial sfGFP solution. In order to prevent non-specific GFP protein adsorption to the plasticware used in the experiment, a sterile solution of 1% BSA (bovine

serum albumin) in PBS was used to block the plastic tubes prior to the experiment. To do this, we filled the 1.5 mL tubes with 1 mL of the BSA solution and left in a rotating mixer for an hour.

2.6.8. *In vitro* study of engineered pMUT plasmid retention and protein expression

The protocol described below was reviewed and approved by the Harvard Medical Area Standing Committee on Animals (HMA IACUC, Ref. No. IS00000516-3). 25 female 8- to 9-week-old C57BL/6NCrl mice were randomly split into five experimental cohorts: WT pMUTs, pM1, pM1-VHH, pM2 and pM2-VHH (N = 5). Bacterial suspensions were prepared in advance by growing to mid-exponential phase (OD600 of 0.5) at 30°C (shaking at 225 RPM), pelleting the cells, resuspending to OD600 of 10 in PBS supplemented with 20% sucrose and 10% glycerol, and flash-freezing in liquid nitrogen. Aliquots of these bacterial suspensions were stored at -80°C and allowed to thaw immediately preceding daily feeding, in order to maintain consistent bacterial density of the inoculum.

48 hours prior to initial administration of bacteria (day -2), the drinking water was supplemented with 2 g/L carbenicillin (Teknova). Antibiotic-free drinking water was restored 24 hours later (day -1). Starting on day 0, each cohort was fed 50 µL of its respective bacterial suspension by allowing the mice to lap the liquid from a pipette tip (as previously described by Mohawk *et al.*[53]). Bacterial administration was carried out daily from day 0 to day 4. Fecal pellets were collected and weighed daily from day 0 to day 7.

2.6.9. Mice

Female 8- to 9-week-old C57BL/6NCrl mice were obtained from Charles River Laboratories. Mice were housed in sterile vinyl isolators within the Harvard Medical School animal facility, and kept under specific-pathogen-free (SPF) conditions. Both sterile food (JL Rat

and Mouse/Auto 6F 5K67, LabDiet) and water were provided ad libitum. All mice were allowed 1 week to acclimate prior to any experimental procedure. To further minimize impact of living environment on experimental outcomes, mice were randomized between housing isolators at the beginning of the experiment. All experiments were conducted in compliance with the US National Institutes of Health guidelines and approved by the Harvard Medical Area Standing Committee on Animals.

2.6.10. Plasmid retention analysis

Immediately following daily collection of fecal pellets, each sample was homogenized in 1 mL of PBS, serially diluted, and plated in quadruplicate to enumerate colony forming units (CFU). Samples were plated on two types of LB agar plates - 25 µg/mL chloramphenicol-only plates (Cm) and 100 µg/mL carbenicillin + 25 µg/mL chloramphenicol plates (Cm+Carb). While all PBP8-derived strains carried a chromosomal Cm resistance gene, only the engineered pMUT plasmids harbored a Carb resistance marker. Plasmid retention rate was therefore estimated by calculating the Cm+Carb to Cm ratio of sample weight-normalized CFU counts. Following the plating procedure, fecal homogenates were flash-frozen and stored at -80°C for subsequent analysis.

2.6.11. Fecal filtration ELISA

To detect E-tagged curli fibers in fecal samples, a filtration ELISA protocol was adapted from Praveschotinunt *et al.*[39]. Fecal homogenate was centrifuged at 1000g for 1 minute to separate large insoluble material, and the supernatant was transferred onto a 96-well filter plate in triplicate (0.22 µm pore size, Multiscreen-GV, Merck/Millipore Sigma). For each sample, the homogenate volume dispensed was normalized to 1.25 mg of fecal pellet per well. After samples were added to the filter plate, the procedure to detect E-tagged material

proceeded as described above in the curli measurement section. In each assay, the signal was normalized by dividing by the WT pMUTs control, such that the WT pMUTs control corresponded to 100%.

2.7. Acknowledgements

This work was done in collaboration with Dr. Anton Kan, Sivaram Emani, Dr. Pichet Praveschotinunt, and Prof. Neel S. Joshi. This work is grateful for and made use of the facilities of the Harvard Center for Comparative Medicine. pFREE was a gift from Morten Norholm (Addgene plasmid #92050). This work was supported by the National Institutes of Health (1R01DK110770-01A1), NSF Grant 1410751 (Division of Materials Research) and the Wyss Institute for Biologically Inspired Engineering. Adapted with permission from: Kan, A., Gelfat, I., Emani, S., Praveschotinunt, P., & Joshi, N. S. (2021). Plasmid Vectors for in Vivo Selection-Free Use with the Probiotic E. coli Nissle 1917. *ACS Synthetic Biology* **2021** *10* (1), 94-106.

2.8. References

1. U. Sonnenborn: Escherichia coli strain Nissle 1917—from bench to bedside and back: history of a special Escherichia coli strain with probiotic properties. *FEMS Microbiol. Lett.* **363**(19) (2016).

2. Altenhoefer Artur, Oswald Sibylle, Sonnenborn Ulrich, Enders Corinne, Schulze Juergen, Hacker Joerg, and Oelschlaeger Tobias A: The probiotic Escherichia coli strain Nissle 1917 interferes with invasion of human intestinal epithelial cells by different enteroinvasive bacterial pathogens. *FEMS Immunol. Med. Microbiol.* **40**(3), 223 (2006).

3. B. Ou, Y. Yang, W. L. Tham, L. Chen, J. Guo, and G. Zhu: Genetic engineering of probiotic Escherichia coli Nissle 1917 for clinical application. *Appl. Microbiol. Biotechnol.* **100**(20), 8693 (2016).

4. C. B. Kurtz, Y. A. Millet, M. K. Puurunen, M. Perreault, M. R. Charbonneau, V. M. Isabella, J. W. Kotula, E. Antipov, Y. Dagon, W. S. Denney, D. A. Wagner, K. A. West, A. J. Degar, A. M. Brennan, and P. F. Miller: An engineered E. coli Nissle improves hyperammonemia and survival in mice and shows dose-dependent exposure in healthy humans. *Sci. Transl. Med.* **11**(475), eaau7975 (2019).

5. T. Danino, A. Prindle, G. A. Kwong, M. Skalak, H. Li, K. Allen, J. Hasty, and S. N. Bhatia: Programmable probiotics for detection of cancer in urine. *Sci. Transl. Med.* **7**(289), 289ra84 (2015).

6. D. S. Leventhal, A. Sokolovska, N. Li, C. Plescia, S. A. Kolodziej, C. W. Gallant, R. Christmas, J.-R. Gao, M. J. James, A. Abin-Fuentes, M. Momin, C. Bergeron, A. Fisher, P. F. Miller, K. A. West, and J. M. Lora: Immunotherapy with engineered bacteria by targeting the STING pathway for anti-tumor immunity. *Nat. Commun.* **11**(1), 2739 (2020).

7. K. Greenhalgh, K. M. Meyer, K. M. Aagaard, and P. Wilmes: The human gut microbiome in health: establishment and resilience of microbiota over a lifetime. *Environ. Microbiol.* **18**(7), 2103 (2016).

8. D. T. Riglar, T. W. Giessen, M. Baym, S. J. Kerns, M. J. Niederhuber, R. T. Bronson, J. W. Kotula, G. K. Gerber, J. C. Way, and P. A. Silver: Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. *Nat. Biotechnol.* **35**(7), 653 (2017).

9. W. R. Whitaker, E. S. Shepherd, and J. L. Sonnenburg: Tunable Expression Tools Enable Single-Cell Strain Distinction in the Gut Microbiome. *Cell* **169**(3), 538 (2017).

10. D. T. Riglar, D. L. Richmond, L. Potvin-Trottier, A. A. Verdegaal, A. D. Naydich, S. Bakshi, E. Leoncini, L. G. Lyon, J. Paulsson, and P. A. Silver: Bacterial variability in the mammalian gut captured by a single-cell synthetic oscillator. *Nat. Commun.* **10**(1), 1 (2019).

11. A. L. Slusarczyk, A. Lin, and R. Weiss: Foundations for the design and implementation of synthetic genetic circuits. *Nat. Rev. Genet.* **13**(6), 406 (2012).

12. C. M. Theriot, M. J. Koenigsknecht, P. E. Carlson, G. E. Hatton, A. M. Nelson, B. Li, G. B. Huffnagle, J. Z. Li, and V. B. Young: Antibiotic-induced shifts in the mouse gut

microbiome and metabolome increase susceptibility to Clostridium difficile infection. *Nat. Commun.* **5**(1), 3114 (2014).

13. M. Schultz, S. Watzl, T. A. Oelschlaeger, H. C. Rath, C. Göttl, N. Lehn, J. Schölmerich, and H.-J. Linde: Green fluorescent protein for detection of the probiotic microorganism Escherichia coli strain Nissle 1917 (EcN) in vivo. *J. Microbiol. Methods* **61**(3), 389 (2005).

14. H. Yano, M. Shintani, M. Tomita, H. Suzuki, and T. Oshima: Reconsidering plasmid maintenance factors for computational plasmid design. *Comput. Struct. Biotechnol. J.* **17**, 70 (2018).

15. F. Hayes: Toxins-Antitoxins: Plasmid Maintenance, Programmed Cell Death, and Cell Cycle Arrest. *Science* **301**(5639), 1496 (2003).

16. A. J. H. Fedorec, T. Ozdemir, A. Doshi, Y.-K. Ho, L. Rosa, J. Rutter, O. Velazquez, V. B. Pinheiro, T. Danino, and C. P. Barnes: Two New Plasmid Post-segregational Killing Mechanisms for the Implementation of Synthetic Gene Networks in Escherichia coli. *iScience* **14**, 323 (2019).

17. C. W. Kang, H. G. Lim, J. Yang, M. H. Noh, S. W. Seo, and G. Y. Jung: Synthetic auxotrophs for stable and tunable maintenance of plasmid copy number. *Metab. Eng.* **48**, 121 (2018).

18. T. E. Kuhlman and E. C. Cox: Site-specific chromosomal integration of large synthetic constructs. *Nucleic Acids Res.* **38**(6), e92 (2010).

19. K. A. Datsenko and B. L. Wanner: One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci.* **97**(12), 6640 (2000).

20. J. Burian, L. Guller, M. Mačor, and W. W. Kay: Small Cryptic Plasmids of Multiplasmid, ClinicalEscherichia coli. *Plasmid* **37**(1), 2 (1997).

21. M. Feldgarden, S. Golden, H. Wilson, and M. A. Riley: Can phage defence maintain colicin plasmids in Escherichia coli. *Microbiology,* **141**(11), 2977 (1995).

22. G. Blum-Oehler, S. Oswald, K. Eiteljörge, U. Sonnenborn, J. Schulze, W. Kruis, and J. Hacker: Development of strain-specific PCR reactions for the detection of the probiotic Escherichia coli strain Nissle 1917 in fecal samples. *Res. Microbiol.* **154**(1), 59 (2003).

23. J. Hacker, T. Oelschlaeger, S. Oswald, U. Sonnenborn, and H. Proppert: US7993902B2 (9 August 2011).

24. H. S. Zainuddin, Y. Bai, and T. J. Mansell: CRISPR-based curing and analysis of metabolic burden of cryptic plasmids in Escherichia coli Nissle 1917. *Eng. Life Sci.* **19**(6), 478 (2019).

25. J. Hacker, U. Sonnen-Born, J. Schulze, G. Blum-Oehler, J. Malinka, and H. Proppert: US6391631B1 (21 May 2002).

26. M. Reister, K. Hoffmeier, N. Krezdorn, B. Rotter, C. Liang, S. Rund, T. Dandekar, U. Sonnenborn, and T. A. Oelschlaeger: Complete genome sequence of the Gram-negative probiotic Escherichia coli strain Nissle 1917. *J. Biotechnol.* **187**, 106 (2014).

27. T. Tatusova, M. DiCuccio, A. Badretdin, V. Chetvernin, S. Ciufo, and W. Li: Prokaryotic Genome Annotation Pipeline (National Center for Biotechnology Information (US), 2013).

28. M. B. Avison, T. R. Walsh, and P. M. Bennett: pUB6060: A Broad-Host-Range, DNA Polymerase-I-Independent ColE2-like Plasmid. *Plasmid* **45**(2), 88 (2001).

29. U. Sonnenborn and J. Schulze: The non-pathogenic Escherichia coli strain Nissle 1917 – features of a versatile probiotic. *Microb. Ecol. Health Dis.* **21**(3–4), 122 (2009).

30. M. Gotfredsen and K. Gerdes: The Escherichia coli relBE genes belong to a new toxin–antitoxin gene family. *Mol. Microbiol.* **29**(4), 1065 (1998).

31. J. P. Torella, F. Lienert, C. R. Boehm, J.-H. Chen, J. C. Way, and P. A. Silver: Unique nucleotide sequence–guided assembly of repetitive DNA parts for synthetic biology applications. *Nat. Protoc.* **9**(9), 2075 (2014).

32. I. Lauritsen, A. Porse, M. O. A. Sommer, and M. H. H. Nørholm: A versatile one-step CRISPR-Cas9 based approach to plasmid-curing. *Microb. Cell Factories* **16** (2017).

33. A. Gyorgy, J. I. Jiménez, J. Yazbek, H.-H. Huang, H. Chung, R. Weiss, and D. Del Vecchio: Isocost Lines Describe the Cellular Economy of Genetic Circuits. *Biophys. J.* **109**(3), 639 (2015).

34. D. I. Piraner, M. H. Abedi, B. A. Moser, A. Lee-Gosselin, and M. G. Shapiro: Tunable thermal bioswitches for *in vivo* control of microbial therapeutics. *Nat. Chem. Biol.* **13**(1), 75 (2017).

35. R. Hurme, K. D. Berndt, E. Namork, and M. Rhen: DNA Binding Exerted by a Bacterial Gene Regulator with an Extensive Coiled-coil Domain. *J. Biol. Chem.* **271**(21), 12626 (1996).

36. V. S. A. Salamov and A. Solovyevand: Automatic annotation of microbial genomes and metagenomic sequences. *Metagenomics Its Appl. Agric. Nova Sci. Publ. Hauppauge NY USA* 61 (2011).

37. A. Muheem, F. Shakeel, M. A. Jahangir, M. Anwar, N. Mallick, G. K. Jain, M. H. Warsi, and F. J. Ahmad: A review on the strategies for oral delivery of proteins and peptides and their clinical perspectives. *Saudi Pharm. J.* **24**(4), 413 (2016).

38. M. M. Barnhart and M. R. Chapman: Curli Biogenesis and Function. *Annu. Rev. Microbiol.* **60**(1), 131 (2006).

39. P. Praveschotinunt, A. M. Duraj-Thatte, I. Gelfat, F. Bahl, D. B. Chou, and N. S. Joshi: Engineered E. coli Nissle 1917 for the delivery of matrix-tethered therapeutic domains to the gut. *Nat. Commun.* **10**(1), 1 (2019).

40. P. Praveschotinunt, N.-M. Dorval Courchesne, I. den Hartog, C. Lu, J. J. Kim, P. Q. Nguyen, and N. S. Joshi: Tracking of Engineered Bacteria In Vivo Using Nonstandard Amino Acid Incorporation. *ACS Synth. Biol.* **7**(6), 1640 (2018).

41. M. L. Evans, E. Chorell, J. D. Taylor, J. Åden, A. Götheson, F. Li, M. Koch, L. Sefer, S. J. Matthews, P. Wittung-Stafshede, F. Almqvist, and M. R. Chapman: The Bacterial

Curli System Possesses a Potent and Selective Inhibitor of Amyloid Formation. *Mol. Cell* **57**(3), 445 (2015).

42. U. Rothbauer, K. Zolghadr, S. Tillib, D. Nowak, L. Schermelleh, A. Gahl, N. Backmann, K. Conrath, S. Muyldermans, M. C. Cardoso, and H. Leonhardt: Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nat. Methods* **3**(11), 887 (2006).

43. M. A. Ghahroudi, A. Desmyter, L. Wyns, R. Hamers, and S. Muyldermans: Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Lett.* **414**(3), 521 (1997).

44. A. Kan, D. P. Birnbaum, P. Praveschotinunt, and N. S. Joshi: Congo Red Fluorescence for Rapid In Situ Characterization of Synthetic Curli Systems. *Appl. Environ. Microbiol.* **85**(13), e00434 (2019).

45. T. J. Rudge, F. Federici, P. J. Steiner, A. Kan, and J. Haseloff: Cell shape-driven instability generates self-organised, fractal patterning of cell layers. *ACS Synth. Biol.* **2**(12), 705 (2013).

46. D. M. Heithoff, C. P. Conner, P. C. Hanna, S. M. Julio, U. Hentschel, and M. J. Mahan: Bacterial infection as assessed by in vivo gene expression. *Proc. Natl. Acad. Sci.* **94**(3), 934 (1997).

47. S. K. Collinson, L. Emödy, K. H. Müller, T. J. Trust, and W. W. Kay: Purification and characterization of thin, aggregative fimbriae from Salmonella enteritidis. *J. Bacteriol.* **173**(15), 4773 (1991).

48. R. H. J. van der Linden, L. G. J. Frenken, B. de Geus, M. M. Harmsen, R. C. Ruuls, W. Stok, L. de Ron, S. Wilson, P. Davis, and C. T. Verrips: Comparison of physical chemical properties of llama VHH antibody fragments and mouse monoclonal antibodies. *Biochim. Biophys. Acta BBA - Protein Struct. Mol. Enzymol.* **1431**(1), 37 (1999).

49. K. Saito: [Studies on the habitation of pathogenic Escherichia coli in the intestinal tract of mice. I. Comparative experiments on the habitation of each type of resistant pathogenic Escherichia coli under an administration of streptomycin]. *Paediatr. Jpn.* **65**, 385 (1961).

50. A. Lerner, T. Matthias, and R. Aminov: Potential Effects of Horizontal Gene Exchange in the Human Gut. *Front. Immunol.* **8** (2017).

51. R. U. Sheth, V. Cabral, S. P. Chen, and H. H. Wang: Manipulating Bacterial Communities by in situ Microbiome Engineering. *Trends Genet.* **32**(4), 189 (2016).

52. J. M. Rhodes: The role of Escherichia coli in inflammatory bowel disease. *Gut* **56**(5), 610 (2007).

53. K. L. Mohawk, A. R. Melton-Celsa, T. Zangari, E. E. Carroll, and A. D. O'Brien: Pathogenesis of Escherichia coli O157:H7 strain 86-24 following oral infection of BALB/c mice with an intact commensal flora. *Microb. Pathog.* **48**(3), 131 (2010).

Chapter 3

Single domain antibodies against enteric pathogen virulence factors are active as curli fiber fusions on probiotic *E. coli* Nissle 1917

3.1. Abstract

Enteric infectious agents represent a major cause of global morbidity and mortality. These pathogens, including viruses, bacteria, and eukaryotic microbes, have an impact in all countries but take a particularly heavy toll in low-income countries where diarrheal disease remains a major cause of infant mortality. For example, infections from several enteric microbial pathogens, including *Escherichia coli*, *Shigella* and *Cryptosporidium* species, were highly associated with infant mortality in the Global Enteric Multicenter Study (GEMS). Novel effective and economical therapeutic approaches are sorely needed. Because of their stability and versatility, VHHs, the variable domains of camelid heavy-chain-only antibodies, have potential as components of novel agents to treat or prevent enteric infectious disease. Here, we describe the isolation and characterization of VHHs targeting several EPEC virulence factors: flagellin (Fla), which is required for bacterial motility and promotes colonization; both translocated intimin receptor (Tir) and intimin, which together play key roles in bacterial association with the host; and *E. coli* secreted protein A (EspA), an essential component of the type III secretion system (T3SS). The virulence factor binding VHHs were characterized for their binding and anti-pathogen properties and a select subset were then displayed on the probiotic strain *E. coli* Nissle 1917 (EcN) by genetically fusing them to curli fibers, the major proteinaceous component of *E. coli* biofilms. We demonstrate that EcN-displayed curli-VHHs bind soluble virulence factors (e.g., Shiga toxin), bacterial surface antigens (e.g., *E. coli* Fla and the T3SS of *Shigella flexneri*) and a major surface antigen of the eucaryotic pathogen *Cryptosporidium parvum*. Taken together, these results demonstrate the efficacy of the curli-based pathogen sequestration strategy described herein and contribute to the development of novel VHH-based gut therapeutics.

3.2. Introduction

Enteric pathogens, which include viruses, bacteria, and eukaryotic microbes, represent a major cause of global morbidity and mortality. These pathogens take a particularly heavy toll in low-income countries where diarrheal disease remains a major cause of infant mortality[1,2]. Traditional interventions such as antibiotics and vaccines suffer from limited efficacy, distribution and implementation challenges, and the rise of antimicrobial resistance[3]. Virulence factors have been identified for many important enteric microbes, but conventional measures to prevent or treat diarrheal disease based on these factors have proved difficult to develop so new therapeutic strategies are needed.

One of the leading causes of infant diarrheal disease and associated mortality in low- and middle-income countries is enteropathogenic *Escherichia coli* (EPEC)[2,4-6]. Colonization by EPEC is facilitated by flagella- (Fla-) driven motility to promote penetration of the mucus layer and association with the intestinal epithelium, where the bacterium induces the formation of 'attaching and effacing' (AE) lesions[7,8]. These lesions, which enable epithelial colonization, are characterized by the effacement of microvilli and the induction of filamentous actin 'pedestals' beneath bacteria closely associated with intestinal epithelial cells[9,10]. To generate AE lesions, EPEC utilizes a type III secretion system (T3SS) to translocate the bacterial effector Tir (translocated intimin receptor) into host cells, where it localizes to the plasma membrane and binds to the EPEC surface adhesin intimin. Intimin-mediated clustering of Tir triggers the assembly of filamentous actin beneath bound bacteria. The related pathogen, Shiga toxin-producing enterohemorrhagic *E. coli* (EHEC), a food-borne pathogen which causes systemic illness in high-income regions such as the U.S. and Europe, generates AE lesions by a similar mechanism[9,11-13], as do some veterinary pathogens such as rabbit enteropathogenic *E. coli* (REPEC) and *Citrobacter rodentium*[14-16].

The direct administration of antibodies or antibody fragments has been proposed as a potential treatment for enteric diseases of diverse etiology[17-23]. VHHs, the variable domain of camelid heavy-chain-only antibodies (also known as 'nanobodies'), appear particularly well suited for this application[18,24-28]. Unlike conventional antibodies, VHH antibodies can be efficiently and functionally expressed in *E. coli* thanks to their small size and single-domain structure. Furthermore, VHHs are effectively expressed as fusion proteins with other VHHs, thus potentially enhancing avidity, increasing specificity, enabling binding to multiple targets[29]. Fusion with other functional domains adds further versatility to their use as therapeutic agents. Together, these properties confer the potential to reduce production costs, improve scalability, and enable novel therapeutic applications.

Although VHHs have opened many novel therapeutic avenues, several challenges remain for their implementation as intestinal therapeutics. Despite the inherent stability and robustness of many VHHs[30], the harsh chemical and enzymatic conditions and continuous flow found in the GI tract will promote the degradation and clearance of VHHs before they reach their target. The delivery and stability of functional VHHs in sufficient quantities to the gut environment therefore constitutes a substantial hurdle. Additionally, producing and purifying large amounts of VHHs is likely to be resource- and labor-intensive, effectively limiting the practicality of such approaches, a particularly relevant limitation for implementation in low-income nations, where the enteric disease burden is highest.

Engineered living therapeutics represent an alternative strategy for the localized production and delivery of molecules into the gut. By genetically modifying a suitable nonpathogenic bacterial strain, heterologous proteins of interest can be expressed *in situ*, circumventing the challenges associated with traditional drug delivery strategies[31-33]. The ability to utilize bacteria themselves as a therapeutic agent, bypassing the need for protein

purification, can potentially render engineered living therapeutics inexpensive and scalable. *E. coli* Nissle 1917 (EcN) has emerged in recent years as a leading candidate for such approaches[34,35]. EcN has displayed an excellent track record of safety through decades of use as a probiotic, and has also been shown to reduce the severity of ulcerative colitis symptoms[36] as well as interfere with the pathogenicity of several enteric pathogens[37], in part due to its ability to colonize the human gastrointestinal tract[38,39]. Transient colonization in humans has also been shown using engineered EcN[40]. Notably, EcN and other laboratory *E. coli* strains can express VHHs as demonstrated by several studies[41-43].

In previous work, we used engineered EcN to display modified curli fibers *in vivo*[44]. Curli fibers are the main proteinaceous components of *E. coli* biofilms. By fusing heterologous protein domains to CsgA – the major curli subunit – we were able to construct a cell-anchored mesh of robust amyloid fibers endowed with novel functionalities, ranging from the display of anti-inflammatory peptides[44] to the nucleation of gold nanoparticles[45]. By fusing VHHs to CsgA, we sought to adapt this strategy to enable the binding of enteric pathogens *in situ*, thereby interfering with their pathogen-host interactions and possibly resulting in pathogen elimination. We call this approach "curli-based pathogen sequestration", drawing an analogy to the polymer sequestrants used to remove excess ions from the gut in chronic kidney disease and a handful of other disorders[46,47].

Here, we describe the generation and characterization of novel VHHs against the Fla, Tir, intimin and EspA antigens of several EPEC, REPEC, EHEC, and *Citrobacter* strains. We then fuse a subset of these VHHs, along with several previously described VHHs that bind virulence factors from other enteric pathogens, to CsgA. By expressing these modified curli fibers in EcN and testing their function, we demonstrate the efficacy of the curli-based sequestration approach *in vitro* against several pathogenic *E. coli* virulence factors. Finally,

we showed that EcN producing CsgA-VHH fusions are capable of recognizing two other major enteric pathogens, *Shigella flexneri* and the eukaryotic pathogen *Cryptosporidium parvum*.

3.3. Results

3.3.1. Generation of VHHs that recognize Fla, Tir, intimin, or EspA

With the goal to obtain the VHHs that bind to virulence factors of AE members of the pathogenic *E. coli* family (Figure 3.1), we performed immunized alpacas with virulence factor that were purified from selected enteric pathogens or prepared as recombinant proteins. For anti-flagellar VHHs, we used both purified REPEC or EPEC flagella or purified recombinant FliC proteins from multiple different pathogenic *E. coli* species. Recombinant REPEC EspA, EHEC and *C. rodentium* intimins, and EHEC Tir protein immunogens were used to elicit anti-EspA, anti-intimin and anti-Tir VHHs, respectively. Unlike flagella, the EspA, intimin, and Tir proteins are relatively well conserved, increasing the likelihood to identify VHHs that recognize diverse AE pathogens. After immunizations, phage-displayed VHHs were panned and then screened against the antigens used for immunization, as well as purified flagella or recombinant proteins representing orthologs from related AE pathogens, as described in Materials and Methods. VHH DNA sequences were determined and one or two representative VHHs were selected for soluble protein expression from each different family of related VHHs (i.e. apparently deriving from a common B cell progenitor). The selected VHHs and their binding properties are summarized below in Table 3.1. Enzyme-linked immunosorbent assays (ELISAs) were used to estimate the apparent affinity (EC_{50} values) for their original antigen as well as

homologous targets from other AE pathogens. VHHs varied widely in their binding capacity and cross-specificity (Table 3.1, Supplementary Figure C1). To simplify VHH nomenclature, some of the selected VHHs were assigned a simplified name related to their target antigen, e.g., "αInt-12" (i.e., "anti-intimin 12").

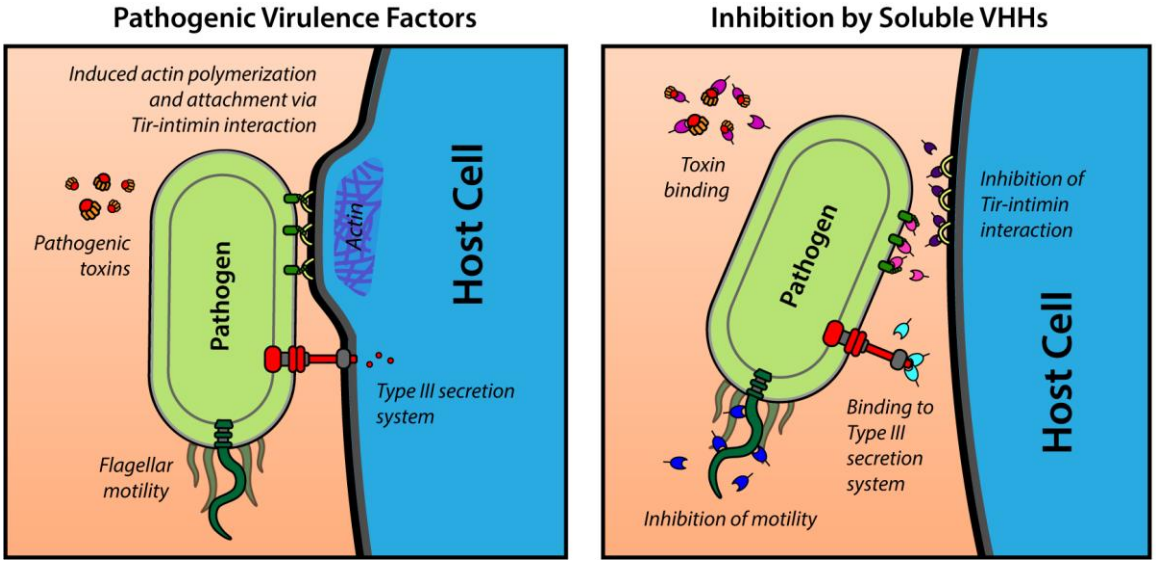


Figure 3.1: Schematic overview of bacterial virulence factors used as VHH targets in this study. Schematic created by Colter Giem.

Despite our success in identifying several VHHs recognizing purified intimin proteins from various pathogenic *E. coli*, each were highly specific to only intimin from one species (not shown) and, importantly, none recognized *E. coli* strain MC1061 expressing EPEC or EHEC intimin on the bacterial surface (e.g., αInt-13, Table 3.1). Since VHHs are particularly dependent on conformational epitopes[48], we hypothesized that the recombinant intimins may not accurately represent the conformation of intimin as displayed by *E. coli*[49,50]. To test this idea, VHH discovery employed *E. coli* K12 strain MC1061 expressing EHEC or EPEC intimin to select and identify phage-displayed VHHs recognizing bacterial surface intimins. This process permitted discovery of VHHs αInt-12, -14 and -17 that recognized both EPEC and EHEC intimin when expressed on the surface of strain MC1061 but did not recognize non-intimin producing MC1061 or recombinant intimin (Table 3.1, Supplementary Figure C1). These data highlight the importance of using, when possible, conformationally native antigens rather than recombinant proteins to identify VHHs of interest.

Table 3.1: Selected VHHs

(a) Anti-Fla VHHs								
VHH name	**Vector name**	**Simplified name**	**Immunogen**	**Panned on**	**EC_{50} [nM] [a]**			**REPEC motility inhibition**
					REPEC flagella	**EPEC flagella**	**EPEC FliC**	
JUV-B11	JVE-2	αFla-1	REPEC flagella; EPEC rFliC	REPEC flagella	Trace [b]	NB	Trace	-
JUV-C4	JVE-4	αFla-2			0.5	1	0.3	ND
JUV-E8	JVE-5	αFla-3			0.5	Trace	NB	+
JUV-G8	JVE-7	αFla-4			5	NB	10	+
JUV-H1	JVE-10	αFla-5			NB	NB	NB	-
JUV-H5	JVE-11	αFla-6			5	NB	25	+
JWU-F3	JXA-1			MC1061/EPEC intimin	10	0.2	ND	ND
JWU-H4	JXA-5				3	10	ND	ND
JXE-B1	JXK-1		EPEC flagella	EPEC flagella	NB	0.1	ND	ND

(b) Anti-EspA VHHs							
VHH name	**Vector name**	**Simplified name**	**Immunogen**	**Panned on**	**EC_{50} [nM] [a]**		**Pedestal blocking activity**
					REPEC EspA	***C. rodentium* EspA**	
JXF-D7	JXM-6	αEspA-1	REPEC EspA	REPEC EspA	0.7	0.7	-
JYB-B1	JYE-1	αEspA-2		JXF-D7-captured REPEC EspA[c]	0.3	0.3	ND
JYB-B8	JYE-2	αEspA-3			0.7	0.7	ND
JYB-D1	JYE-3	αEspA-4			3	3	ND
JYB-H4	JYE-4	αEspA-5			0.5	0.5	ND
JYB-H6	JYE-5	αEspA-6			0.3	0.3	ND
JXF-D8	JXM-8	αEspA-7		REPEC EspA	0.7	0.7	-
JXF-H9	JXM-12	αEspA-8			0.7	0.7	-
JXF-C4	JXM-15	αEspA-9			Trace	0.7	-

Table 3.1: Selected VHHs (continued)

(c) Anti-Tir VHHs								
VHH name	**Vector name**	**Simplified name**	**Immunogen**	**Panned on**	**EC_{50} [nM] [a]**			**Tir-intimin blocking activity [d]**
					EHEC Tir	EPEC Tir	REPEC Tir	
JVB-C6	JVG-1	αTir-1	EHEC Tir	REPEC Tir	trace	50	trace	-
JVB-G4	JVG-2	αTir-2			0.1	0.1	0.1	++ [e]
JVB-G8	JVG-3	αTir-3			0.2	0.2	0.2	-
JVC-C6	JVI-1	αTir-4		EHEC Tir	0.1	0.2	0.1	+
JVC-D10	JVI-2	αTir-5			0.1	0.2	0.2	-
JVC-E5	JVI-3	αTir-6			0.1	0.2	0.2	+
JVA-A1	JVF-1	αTir-7			0.7	10	3	-
JVA-C8	JVF-2	αTir-8			0.1	0.2	0.2	+++
JVA-C9	JVF-3	αTir-9			0.1	5	0.2	-
JVA-D4	JVF-4	αTir-10			0.5	0.5	0.5	-
JVA-F6	JVF-7	αTir-11			0.1	25	0.2	-
JVA-D11	JVF-8	αTir-12			0.2	0.2	0.2	+
JVA-E10	JVF-12	αTir-13			0.5	0.2	0.1	-
JVA-G1	JVF-14	αTir-14			0.1	0.2	0.2	+++

(d) Anti-intimin VHHs								
VHH name	**Vector name**	**Simplified name**	**Immunogen**	**Panned on**	**EC_{50} [nM] [a]**			**Pedestal blocking activity [f]**
					EHEC intimin	**EPEC intimin**	**MC1061 /EPEC intimin**	
JWS-H4	JWZ-5	αInt-12	EHEC intimin	*E. coli* 1061/pInt (EHEC)	Trace	Trace	10	-
JWT-C1	JWZ-7	αInt-13			0.5	Trace	NB	+
JWU-D8	JWZ-9	αInt-14		*E. coli* 1061/pInt (EPEC)	Trace	Trace	0.5	+
JWU-G8	JWZ-15	αInt-17			NB	NB	0.5	+
JXN-E2[e]	JXS-2		*C. rodentium* intimin	DH5α/pInt (*C. rodentium*)	NB	NB	NB	ND

[a] EC_{50} estimates based on dilution ELISAs such as shown in Supplementary Figure C1
[b] Trace – EC_{50} >125 nM, i.e., poor but detectable
[c] Panning employed JXF-D7-captured REPEC EspA target
[d] Tir-intimin binding inhibition from Figure 2: + $p<0.01$; ++ $p<0.001$; +++ $p<0.0001$
[e] αTir-2 also displayed pedestal blocking activity, Figure 3.3
[f] Pedestal blocking activity from Figure 3.3
[g] JXN-E2 binds to *C. rodentium* intimin with EC_{50} ~0.5 nM, not shown
NB – no binding; ND – not done

3.3.2. Anti-Fla VHHs inhibit REPEC motility

Flagellin (FliC) is the major protein of flagella which are required for motility and colonization of some pathogenic *E. coli*[51]. We found that REPEC was highly motile when spotted onto low percentage agar plates (Figure 3.2a) and utilized this strain to test whether various anti-Fla VHHs could impair motility. REPEC was grown to mid-log phase and incubated with either PBS or different concentrations of VHHs (as described in Materials and Methods). After depositing bacteria onto the center of an agar plate, REPEC motility was assessed by measuring the growth diameter. Figure 3.2a shows results for αFla-1 VHH which did not inhibit motility, and αFla-6 which did inhibit motility. Two other Fla VHHs, αFla-3 and 4 also showed marked inhibition of REPEC motility at 6.1 and 2.4 μM concentrations (Figure 3.2b), with an IC_{50} as low as 1.0 μM (Table 3.1), whereas αFla-5 showed no motility inhibition, even at 6.1 μM (Figure 3.2b).

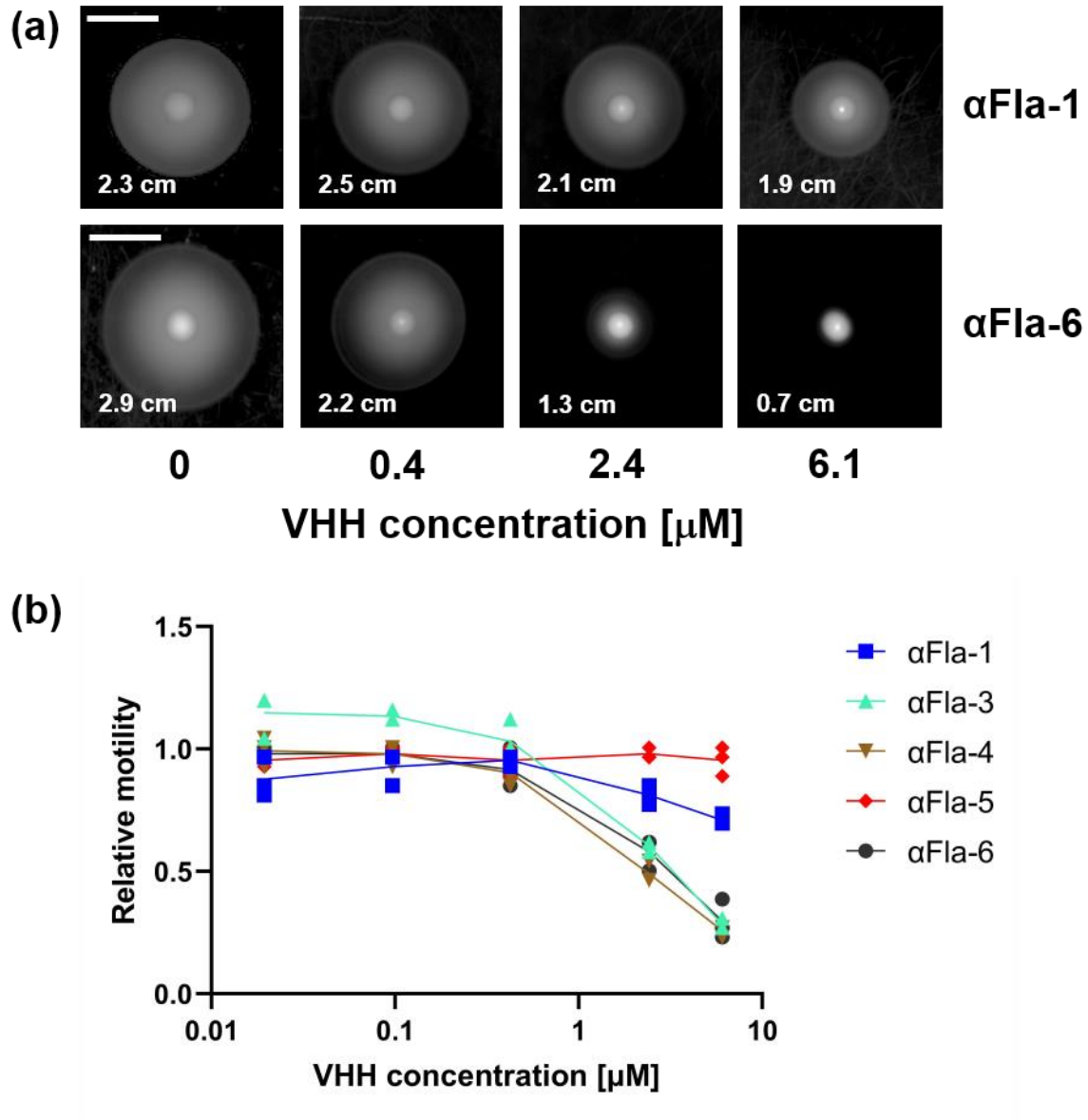


Figure 3.2: Anti-Fla VHHs can inhibit REPEC motility. (a) Representative images of REPEC growth on motility plates after incubation with varying VHH concentrations (scale bar = 1 cm). (b) Relative motility as a function of αFla VHH concentration. Spread diameters were normalized to a "no VHH" control.

3.3.3. Anti-Tir and anti-intimin VHHs that block the interaction between intimin and Tir

The interaction between intimin and Tir is required for bacterial infection of the host. In order to determine if anti-Tir VHHs can block this interaction, ELISA-based assays were performed in which plates were coated with GST-tagged EPEC Tir then incubated with VHHs or (as a negative control) 0.1% BSA. Wells were incubated with either GST-tagged EPEC intimin, or with (as a negative control) GST alone, and bound GST or GST-intimin was detected with anti-GST antibody. As shown in Figure 3.3a, 0.1% BSA showed no ability to block the intimin-Tir interaction, whereas six out of fourteen anti-Tir VHHs showed significant neutralization activity with αTir-2, αTir-8 and αTir-14 displaying the most potent activity. Thus, despite high affinity for soluble Tir, many anti-Tir VHHs were ineffective in blocking the interaction between purified Tir and intimin *in vitro*.

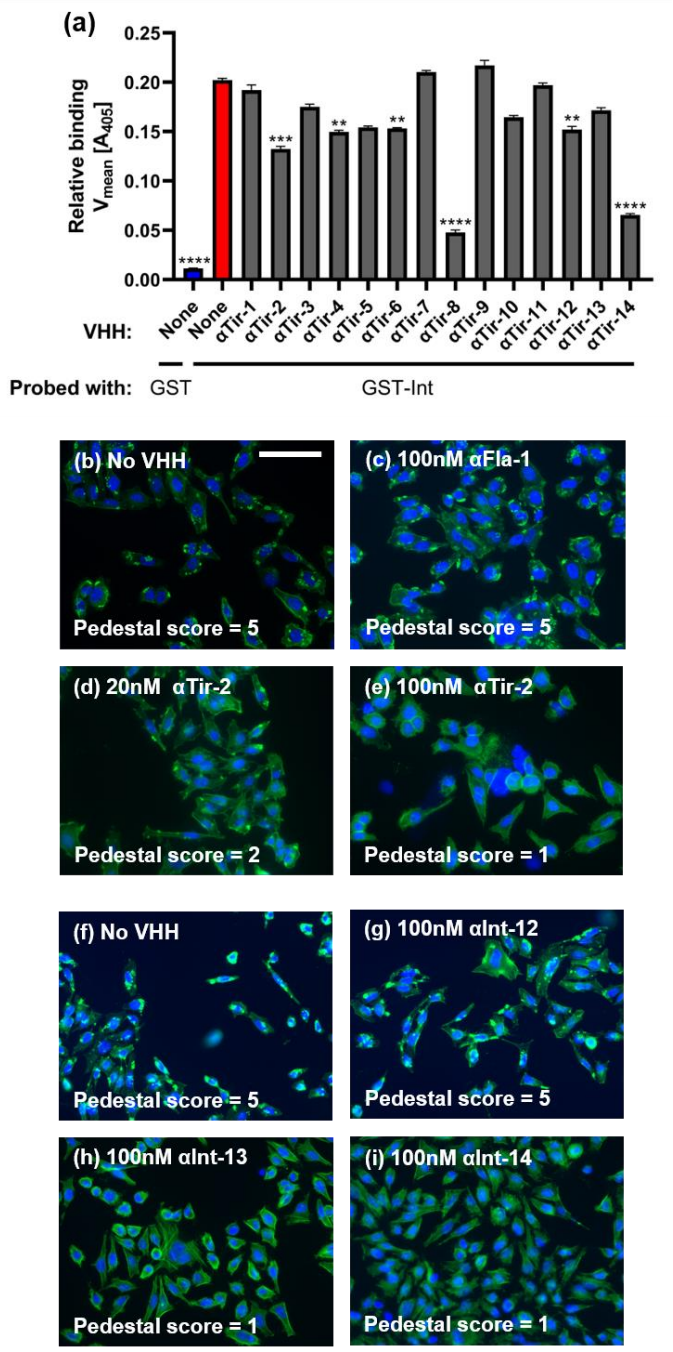


Figure 3.3: Anti-Tir and anti-intimin VHHs interfere with Tir-intimin binding and pedestal formation. (a) Several anti-Tir VHHs blocked Tir-intimin binding. Data presented as mean ± SEM. P-values calculated by one-way ANOVA. ** $P < 0.01$; *** $P < 0.001$, **** $P < 0.0001$. (b-i) Anti-Tir and anti-intimin VHHs can inhibit EPEC-mediated pedestal formation (scale bar = 100 µm). HeLa cells were exposed to EPEC incubated with VHH, fixed and stained with DAPI (blue) and Alexa Fluor-488 Phalloidin (green). Blinded pedestal scores were assigned by a separate researcher. Assays performed by Yousuf Aqeel, Anishma Shrestha, and Shenglan Hu.

3.3.4. EPEC pedestal formation is inhibited by anti-Tir and anti-intimin VHHs

During EPEC infection, bacterial Tir protein on host cells binds to bacterial surface intimin and this clustering promotes F-actin assembly beneath bacteria bound at the host plasma membrane and results in the formation of actin pedestals that facilitate pathogen colonization. As αTir-2 was shown to block the intimin-Tir interaction, we tested whether it was also able to inhibit EPEC pedestal formation. HeLa cells were incubated with EPEC and 100 nM αTir-2 VHH for 3 h at 37°C, then stained with DAPI (to stain cell nuclei) and phalloidin (to stain F-actin). As expected, EPEC infection in the absence of αTir-2 VHH resulted in robust pedestal formation (Figure 3.3b), and pedestal formation was not inhibited by αFla-2, a negative control (Figure 3.3c). Coded images of each sample were scored blindly on a scale of 1 to 5 (see Materials and Methods) and each yielded a pedestal score of 5 (many pedestals). However, the addition of 20 nM αTir-2 partially inhibited pedestal formation (Figure 3.3d) and 100 nM αTir-2 completely inhibited EPEC pedestal formation with no observable pedestal formation (Figure 3.3e). Thus, αTir-2 was shown *in vitro* to inhibit both the interaction between recombinant Tir and intimin, and EPEC pedestal formation, which is triggered by the interaction of these two virulence factors.

Anti-intimin VHHs were also tested for the ability to inhibit EPEC pedestal formation. While αInt-12 showed no inhibition of pedestal formation (Figures 3.3f and 3.3g), αInt-13 and -14 both displayed inhibitory activity (Figures 3.3h and 3.3i). In addition, αInt-17, which is closely related to αInt-14, also showed potent inhibitory effects, although unblinded scoring not performed (Table 3.1, data not shown).

Finally, we tested αEspA-1, 7, 8, and 9 for their ability to block pedestal formation (αEspA-1, 7 and 8 are closely related). Despite their high-affinity binding, none of these anti-EspA VHHs were able to inhibit pedestal formation (Supplementary Figure C2). However, a

different panel of anti-EspA VHHs recognizing a different, non-competing epitope were later found by panning the library on αEspA-1 VHH-captured EspA (JYB series, Table 3.1, Supplementary Table C2) which blocked the apparent immunodominant epitope, but these were not tested for neutralization of pedestal formation.

3.3.5. Design and expression of curli-VHH fusions in EcN

Next, we selected several anti-virulence factor VHHs with which to construct CsgA fusions, with the aim of displaying these curli-tethered binding domains on the surface of EcN. Twelve VHH sequences described above or previously were selected as CsgA fusion candidates (Table 3.2). The CsgA and VHH coding sequences were fused in-frame and inserted into expression plasmid pL6FO, in which a synthetic version of the full curli operon (*csgBACEFG*) was placed under control of an inducible promoter (Figure 3.4b). The two protein domains were connected by a 14 amino acid glycine-serine linker, and a 6xHis-tag was appended to the C-terminus of the VHH domain to facilitate detection. To eliminate potential confounding effects of native CsgA, the plasmids were introduced into EcN strain PBP8 in which the native curli operon was deleted from the genome[44,52].

Following transformation and induction, curli expression and fiber formation were assessed by a Congo Red binding assay, which detects curli on the surface of bacteria[45,53]. Although the degree of Congo Red binding varied slightly between VHH fusions, overall signal remained high compared to the negative control condition, i.e., plasmid-free strain PBP8 containing neither the wild-type *csgA* gene nor any curli fusion (Figure 3.4e). This result therefore indicates the formation of cell-anchored amyloid fibers on the CsgA-VHH-expressing strains. To further characterize the fusion proteins, field emission scanning electron microscopy images of CsgA-VHH fibers were captured. All CsgA-VHH constructs tested resulted in the formation of curli material (Figure 3.4d, Supplementary Figure C3).

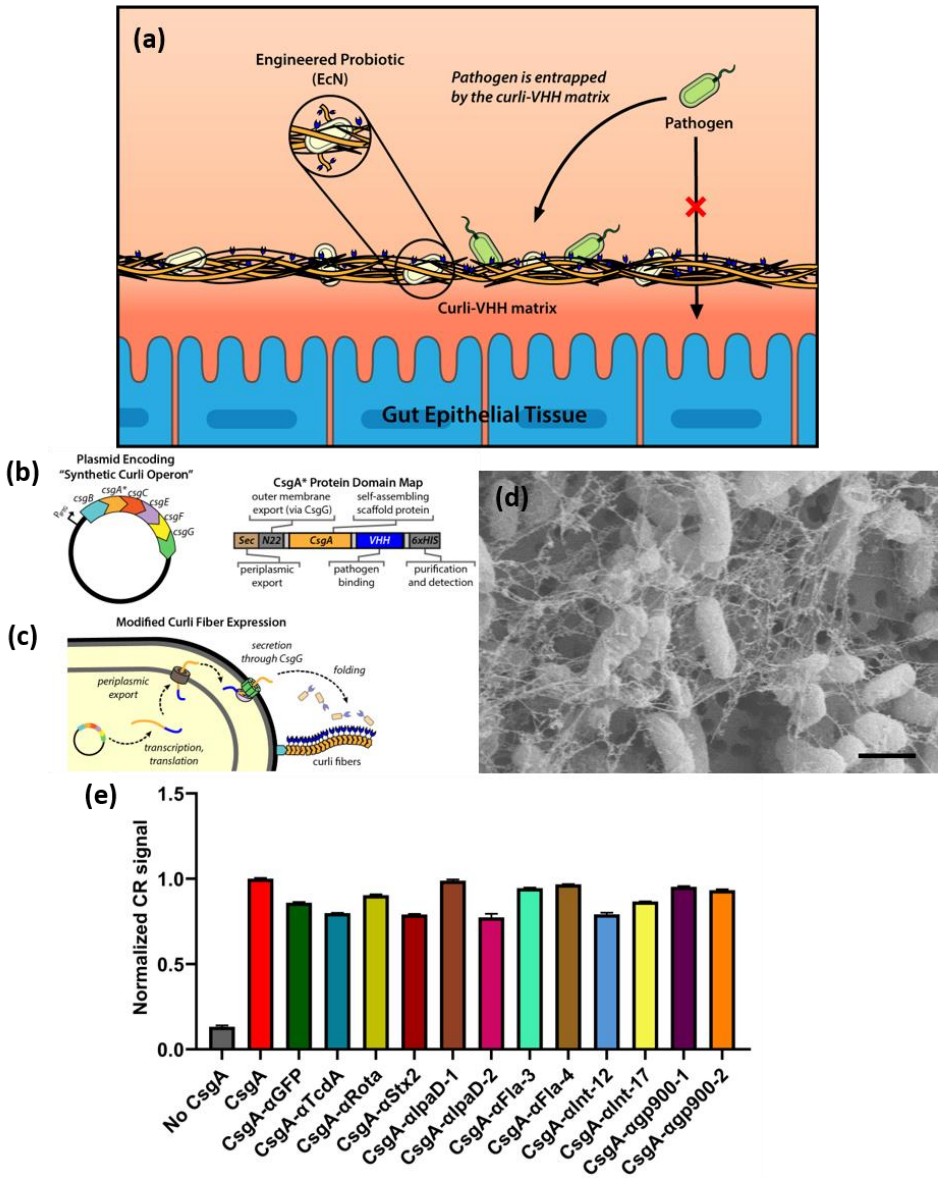


Figure 3.4: Formation of VHH-functionalized curli fibers by EcN. (a) Schematic overview of curli-based pathogen sequestration approach. (b) Plasmid map of synthetic curli operon and domain map of CsgA-VHH fusion constructs. (c) Schematic of curli fiber export and self-assembly mechanism. (d) Representative SEM image demonstrating the formation of curli fibers by CsgA-VHH-expressing EcN (CsgA-αgp900-2, scale bar = 1 µm). (e) Congo Red (CR) binding assays of induced cultures of PBP8 bearing plasmids encoding various CsgA-VHH constructs. CR binding is indicative of the formation of curli amyloid fibers. Data corresponds to A_{490} measurements normalized to the CsgA positive control, and is presented as mean ± SEM. Two-way ANOVA ($P < 0.0001$) was performed to test presence of difference between conditions. All CsgA constructs exhibited P-values < 0.0001 compared to the "No CsgA" control, as calculated by Welch's t-test. Schematics designed by Colter Giem and Prof. Neel Joshi. Assays performed by Ilia Gelfat, SEM images captured by Avinash Manjula-Basavanna.

Table 3.2: CsgA-VHH constructs

Construct Name	VHH Name	Target	VHH Source
CsgA	N/A	N/A	N/A
CsgA-αGFP	NbGFP	GFP	Rothbauer *et al.* [54]
CsgA-αStx2	JGH-G1	Shiga toxin 2	Tremblay *et al.* [25]
CsgA-αTcdA	NbTcdA	*C. difficile* toxin TcdA	Hussack *et al.* [55]
CsgA-αRota	3B2	Rotavirus inner capsid protein VP6	Vega *et al.* [26]
CsgA-αIpaD-1	20ipaD	*S. flexneri* T3SS	Barta *et al.* [20]
CsgA-αIpaD-2	JPS-G3	*S. flexneri* T3SS	Barta *et al.* [20]
CsgA-αFla-3	JUV-E8 (αFla-3)	REPEC flagellin	This study
CsgA-αFla-4	JUV-G8 (αFla-4)	REPEC flagellin	This study
CsgA-αInt-12	JWS-H4 (αInt-12)	EPEC intimin	This study
CsgA-αInt-17	JWU-G8 (αInt-14)	EPEC intimin	This study
CsgA-αgp900-1	JJ-D1	*C. parvum* antigen gp900	Jaskiewicz *et al.* [56]
CsgA-αgp900-2	JMP-F7	*C. parvum* antigen gp900	Jaskiewicz *et al.* [56]

3.3.6. CsgA-VHH-producing EcN can bind soluble protein targets and neutralize a potent bacterial toxin

Having demonstrated the formation of curli fibers on EcN from engineered CsgA-VHH monomers, we next tested the functionality of the fused VHH domains. To assess the ability of CsgA-αGFP to target GFP, curli-VHH-producing EcN were cultured and incubated with GFP suspended in simulated colonic fluid. After 15 minutes of incubation at 37°C, the suspensions were pelleted, and the fluorescence of the supernatant was measured to quantify the remaining unbound GFP by comparing the signal to a calibration curve of

known GFP concentrations. The unbound GFP concentration was then subtracted from the initial solution concentration to estimate bound GFP, as shown in Figure 3.5a. No significant binding was observed in the EcN pellet when the unfused CsgA control or an off-target (anti-rotavirus) VHH was used. In contrast, upon pelleting EcN expressing CsgA-αGFP, the signal localized to the pellet was at least 56-fold higher than the negative controls, with the GFP concentration in the supernatant decreased significantly (Figure 3.5a).

To determine if curli-VHH fibers can sequester a potent bacterial toxin from solution, we tested EcN producing a previously characterized VHH directed against Shiga toxin 2 (Stx2)[25]. Stxs are potent bacterial toxins comprising several variants produced by the AE pathogen EHEC, as well as by *Shigella dysenteriae*, and are responsible for the life-threatening manifestations of infections by these pathogens[57]. EcN with curli fiber-displayed VHHs were suspended in 10 ng/mL purified Stx2 and incubated for 1 hour at 37°C. After pelleting the EcN, monolayers of Vero cells were treated with the supernatant for 48 hours and their viability measured by a PrestoBlue cell viability assay. Supernatants derived from control EcN producing no CsgA, CsgA alone, or CsgA fused to either of two VHH directed against irrelevant proteins (GFP or TcdA, an unrelated toxin) were toxic, yielding ~20% Vero cell viability. In contrast, incubation of Stx2 with CsgA-αStx2-producing EcN significantly reduced Stx2-induced toxicity, as ~75% of Vero cells survived treatment with supernatant ($P < 0.0001$; Figure 3.5b).

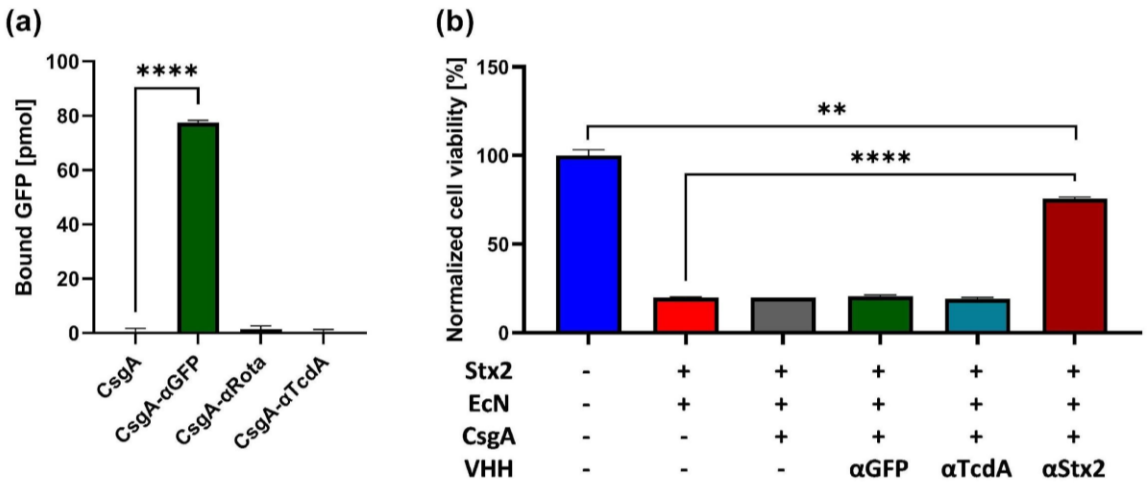


Figure 3.5: CsgA-VHHs can bind soluble proteins. (a) GFP pull-down assay. After incubation with 150 nM GFP, EcN CsgA-αGFP bound specifically to its target, as demonstrated by the depletion GFP in the cell supernatant. (b) Stx2 pull-down assay. EcN CsgA-αStx2 was used to selectively remove Stx2 upon incubation in a 10 ng/mL Stx2 solution. Supernatants were added to Vero cell monolayers and cell viability was measured. Data presented as mean ± SEM. Two-way ANOVA ($P < 0.0001$) was performed to test presence of difference between conditions, P-values calculated by Welch's t-test. ** $P < 0.01$; **** $P < 0.0001$. Assays performed by Ilia Gelfat (a) and James N. Lee (b).

3.3.7. CsgA-VHH-producing EcN can bind to targets on the bacterial cell surface of enteric pathogens

To determine if CsgA-VHH-producing EcN can neutralize virulence factors physically associated with pathogens, we used VHH sequences against several cell-anchored targets of multiple microorganisms and tested their ability to exhibit binding and/or mitigate virulence. We first generated CsgA-VHH constructs targeting the flagellar proteins of REPEC employing VHHs αFla-3 and 4, which had been shown to inhibit REPEC motility (Figure 3.2). To test the ability of these engineered EcN strains to bind REPEC, an aggregation assay was used. When planktonic bacterial cells are added to a conical-well, they gradually settle to the central point at the bottom of a conical well and can be visualized as a focused pellet. In contrast, in the presence of an aggregant, cells instead

form a lattice that blankets the bottom of the well uniformly. Hence, visualization of diffusely distributed cells instead of a focused pellet indicates microbial aggregation. Because the efficiency of aggregation is subject to the ratio of aggregant to cells, we mixed suspensions of REPEC and EcN transformants at different concentrations, then photographed the plates after settling for a day (Supplementary Figure C4). Aggregation was observed in a concentration- and VHH-dependent manner, demonstrating specific binding of the anti-Fla curli-VHHs to their targets while no aggregation occurred testing EcN expressing control VHHs. Specifically, EcN expressing CsgA-αFla-3 or -4 were able to trigger more pronounced aggregation as their concentration increased from 0.1x compared to their overnight culture density to 0.3x, 1x, and 3x. As a target control, the same EcN suspensions were mixed with rabbit enterohemorrhagic *E. coli* (REHEC), which is related to REPEC but produces an antigenically distinct Fla, and no aggregation was observed (Supplementary Figure C4).

Flagella, which extend 5-20 μm from the surface of *E. coli*, should be more easily bound by curli-displayed VHHs than structures closely associated with the microbial surface where the binding may be sterically constrained[58]. Type III secretion systems (T3SSs) extend less than 50 nm from the outer membrane of Gram-negative bacteria[59]. We next tested EcN expressing VHHs targeting a T3SS virulence factor for binding and pathogen neutralization.

Shigella flexneri, a significant contributor to worldwide diarrheal disease burden[2], encodes a T3SS that is essential for virulence. IpaD, which assembles at the distal end of the T3SS apparatus, prevents premature exposure of the effectors to the extracellular environment[60,61]. IpaD has been shown to be involved in pore formation in the host membrane by regulating IpaB and IpaC, and inhibiting IpaD function has been shown to

reduce *S. flexneri*'s ability to disrupt host cell[62]. This pore formation process is typically evaluated using a contact-mediated hemolysis assay measuring red blood cell lysis following exposure to *S. flexneri*[63]. Two anti-IpaD VHHs (αIpaD-1 and αIpaD-2, Table 3.2), previously shown to inhibit red blood cell lysis by *S. flexneri* when applied as soluble proteins[20], were expressed on EcN curli. Both curli-VHH constructs were shown capable of binding soluble IpaD, though CsgA-αIpaD-2 exhibited a much stronger signal than CsgA-αIpaD-1 as measured by ELISA (Supplementary Figure C5).

The ability of these curli-VHHs to neutralize *S. flexneri* was next tested in a hemolysis assay. Suspensions of *S. flexneri* were incubated with EcN expressing either CsgA-αIpaD-1, CsgA-αIpaD-2, or CsgA-αGFP as a negative control. As a positive control, a soluble trimer of αIpaD VHHs was prepared and tested (Figure 3.6). After incubation of EcN and *Shigella*, the bacterial suspensions were exposed to sheep blood, and hemolysis was measured colorimetrically. EcN producing CsgA-αIpaD-2 abolished much of the observed hemolysis as did the positive control. Surprisingly, CsgA-αIpaD-1 did not achieve similar results, despite its ability to bind soluble IpaD (Figure 3.6), suggesting this VHH was sterically inhibited from binding *Shigella* TS33 when expressed in curli fibers.

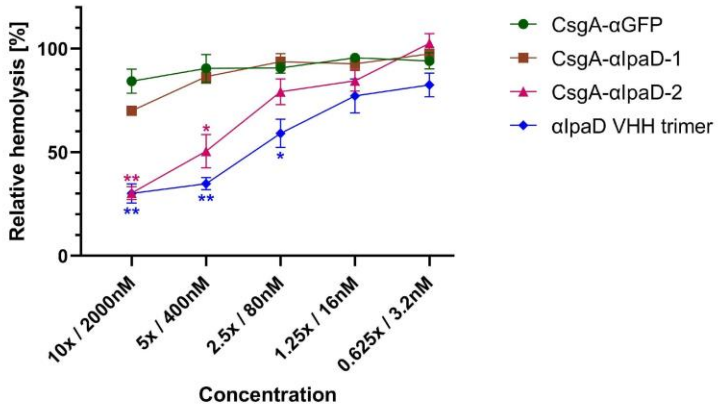


Figure 3.6: CsgA-VHHs can neutralize *S. flexneri* contact-mediated hemolysis. EcN expressing CsgA-αIpaD-2, but not CsgA-αIpaD-1, inhibited hemolysis of sheep red blood cells, significantly outperforming the off-target negative control (CsgA-αGFP). Data presented as mean ± SEM. Two-way ANOVA ($P < 0.0001$) was performed to test the presence of difference between conditions, P-values calculated by Welch's t-test. * $P < 0.05$; ** $P < 0.01$. Assay performed by Ilia Gelfat.

3.3.8. CsgA-VHH-producing EcN can bind to the eukaryotic pathogen *Cryptosporidium parvum*

To determine if CsgA-VHH-producing EcN can bind to a surface expressed protein target on a eukaryotic pathogen, we utilized the parasite *Cryptosporidium parvum*, which, along with *C. hominis*, is the major cause of cryptosporidiosis. Cryptosporidiosis is an enteric diarrheal disease that is a major cause of morbidity in children in low-income countries[1,2], and is the leading cause of waterborne disease in the United States, with prevalence of infections continuing to rise[64]. Cryptosporidiosis can also be severe in immunocompromised individuals such as AIDS/HIV patients, where the prevalence of enteric protozoan infection was reported to be 30.6%[65]. Novel treatments for cryptosporidiosis are urgently needed, as the only available treatment is nitazoxanide, which has limited efficacy and is effective in only a subset of patients[66].

Because of the difficulty of studying and propagating human pathogenic *Cryptosporidium* spp., there are relatively few documented virulence factors or well-characterized surface-exposed antigens compared to many other enteric pathogens. One such antigen, glycoprotein gp900, is expressed on the surface of *C. parvum* sporozoites and has been implicated in motility and host cell invasion[67]. Importantly, gp900 is also shed in trails by sporozoites as they move about.

Previous studies identified VHHs that recognize the carboxyl-terminal domain of gp900 (αgp900-1 and αgp900-2, Table 3.2)[56]. We used both VHH sequences to generate CsgA fusions, expressed these constructs in EcN and tested their ability to bind their target antigen. We first tested gp900 binding using an ELISA, where CsgA-VHH-expressing EcN were adsorbed onto plastic and incubated with *C. parvum* lysate. EcN that expressed CsgA-αgp900-2, though not CsgA-αgp900-1, produced a weak binding signal to indicate binding at high concentrations, which differed significantly from the CsgA-αGFP negative control (Figure 3.7a). Similarly, in a pull-down assay, EcN expressing CsgA-αgp900-2, but not αgp900-1, depleted gp900 from *C. parvum* lysate (Figure 3.7b).

Binding to the parasites was then assayed by fixing *C. parvum* sporozoites onto slides, exposing them to EcN, quantifying the parasite-bound EcN. Interestingly, while CsgA-αgp900-1 did not appear to bind the antigen in its soluble form, it significantly outperformed CsgA-αgp900-2 in pathogen binding, although CsgA-αgp900-2 also bound the parasites significantly better than the CsgA-αGFP negative control (Figure 3.7c). Qualitatively, colocalization of EcN CsgA-αgp900 with the fixed sporozoites was also observed (Supplementary Figure C6).

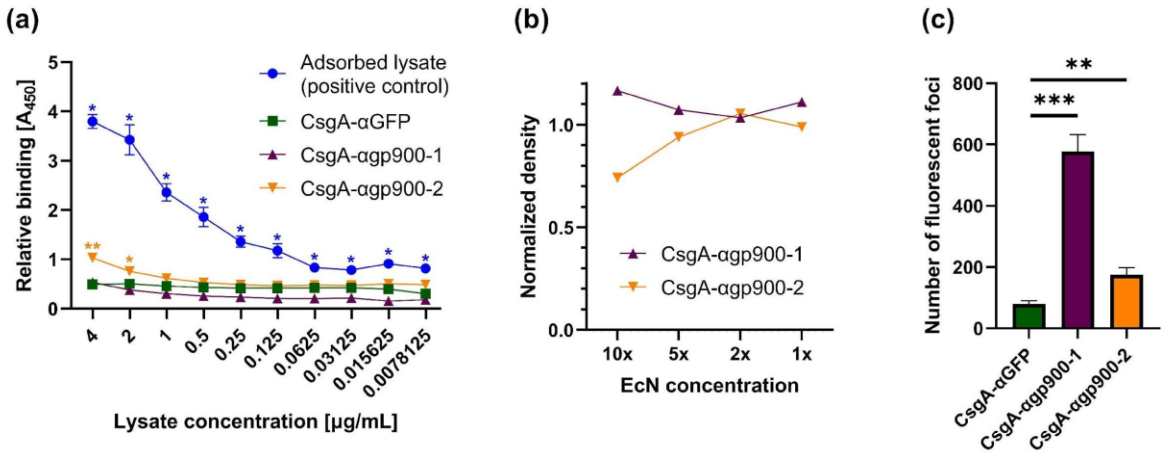


Figure 3.7: CsgA-VHHs can bind to the eukaryotic pathogen *Cryptosporidium parvum*. (a) ELISA demonstrating the ability of CsgA-αgp900-2 to bind gp900. EcN was adsorbed onto a well plate, followed by incubation with gp900-containing *C. parvum* lysate. In the positive control, no EcN was used, and the *C. parvum* lysate was allowed to directly adsorb onto the surface of the wells. Gp900, either bound to EcN or adsorbed to the surface, was detected using a non-competing anti-gp900 VHH, followed by an anti-Etag IgG-HRP conjugate. (b) Western blot analysis of gp900 binding to CsgA-αgp900-2. *C. parvum* lysate was incubated with EcN at different concentrations. After pelleting the EcN, supernatant form was run on gel and subjected to Western blot to assess gp900 depletion. The gp900 band was detected by a specific non-competing VHH, followed by an anti-Etag IgG-HRP conjugate. Band intensity was normalized to a CsgA-αGFP negative control. (c) CsgA-αgp900-1 and -2 bind to *C. parvum* sporozoites. EcN was applied to immobilized and fixed sporozoites, followed by staining with anti-LPS Mab and anti-mouse IgG Alexa Fluor 568. Slides were inspected under TRITC filter and 5 images were taken under 200x magnification for each condition. Foci of EcN accumulation were quantified using ImageJ particle analyzer. Data presented as mean ± SEM. Two-way ANOVA ($P < 0.0001$) was performed to test presence of difference between conditions, P-values calculated by Welch's t-test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. Assays performed by Justyna Jaskiewicz.

3.4. Discussion

In this report, we propose a novel approach, which we term 'curli-based pathogen sequestration', to treatment or prevention of enteric infectious disease. By appending VHH domains to curli fibers and displaying these modified fibers on the probiotic strain EcN, we demonstrated binding, and in some cases – neutralization, of virulence factors of several

pathogens. These factors included the potent bacterial toxin Stx2, the flagella of REPEC, the T3SS of *S. flexneri*, and the surface antigen gp900 of the eukaryotic parasite *C. parvum*. In addition, we described the identification and characterization of VHHs that bind to various surface-exposed virulence factors of EPEC and EHEC. Several of these VHHs inhibited flagella-driven motility or interfered with Tir-intimin binding, and exhibited neutralizing activity against their target pathogen. Adding to the growing body of validated VHH sequences, these antibody fragments may find use in myriad therapeutic, diagnostic, and research applications, and contribute to the study and treatment of pathogenic *E. coli*.

We propose here the further exploration of EcN curli fiber-displayed VHHs as a potential therapeutic approach to the prophylactic prevention of enteric pathogen establishment. We suggest that by maintaining a level of selected CsgA-VHH-producing EcN in the GI tract, the probiotic bacteria may sequester incoming pathogens before they reach their target sites of infection, allowing their elimination from the body by natural processes without causing pathology. Employing EcN-displayed VHHs that bind virulence factors and neutralize the ability of the pathogen to infect the host could add additional protection. The efficacy of curli-based sequestration has not yet been tested *in vivo*, and requires further validation. Specifically, such validation will necessitate the development of animal infection models that support coadministration and maintenance of pathogen and EcN at stable levels, allowing for testing of the efficacy of the curli-based sequestration matrix. Nevertheless, the approach has several features that make it particularly well-suited for binding enteric pathogens *in situ*. First, the curli matrix can provide multivalent display of binding sites, as each CsgA monomer is linked to a VHH domain. Multivalency has been shown to be an important factor for the performance of several pathogen binding systems[25,68-71]. Secondly, curli-based materials have been demonstrated to work in the

gastrointestinal tract, both in the form of anti-inflammatory peptides displayed on curli fibers produced by EcN[44], as well as similarly functionalized curli hydrogels[72].

Additionally, because EcN can replicate in the host and continuously produce new curli fibers, we anticipate that steady-state levels of the multimeric displayed VHHs can be maintained for significant periods of time despite the harsh proteolytic environment and constant flow. Notably, EcN's ability to maintain a steady-state density in the mammalian GI tract with regular dosing is supported by data in mice and humans[35,40]. EcN has also made progress as an engineerable chassis organism for therapeutic applications in the gut in clinical trials[73].

Lastly, the use of engineered microbes to deliver and produce the sequestrant *in situ* may offer additional benefits. Unlike inert biomaterials, bacteria can sense their environment and respond to changes within it. As such, engineered living therapeutics can exert additional therapeutic or diagnostic functions in tandem with the production of pathogen-binding molecules – either in the form of genetic circuits or through their inherent native properties. In particular, the use of EcN as a chassis organism may prove advantageous in targeting EPEC and EHEC variants, as it has been shown to outcompete these strains in biofilm formation[38], prevent EHEC colonization in mice by occupying a similar nutritional niche[74], as well as promote intestinal health through several other mechanisms[75]. Therefore, while further studies are needed, we hypothesize that the pathogen sequestration strategy may work synergistically with EcN's probiotic functions, potentially resulting in a more effective treatment compared to either wild-type EcN or a non-living curli-VHH material.

Notably, not all CsgA-VHH constructs proved equally efficacious. For instance, while both anti-IpaD VHHs tested were shown to exhibit neutralizing activity in their soluble form[20],

only CsgA-αIpaD-2 retained this ability when displayed on EcN-bound curli fibers. This may be due to steric effects related to the spatial arrangement of neighboring VHHs fused to CsgA domains, or may be related to CsgA-VHH secretion and curli fiber assembly, as translocation of some VHH domains through the curli secretion machinery may interfere with proper protein folding. The curli secretion and assembly pathway involves multiple steps and requires translocation through the inner membrane into the periplasm via the SecYEG translocon, followed by transit through the dedicated CsgG outer membrane pore[76]. While this pathway has been shown to accommodate various CsgA fusions, the complete scope of specific limitations on this capacity are yet to be fully elucidated[77]. Nevertheless, its impressive amenability to a wide variety of VHH fusions demonstrates the modularity and flexibility of functionalized curli in general, and VHH-based applications specifically.

In recent years, other pathogen sequestration approaches have been explored for the binding and removal of viruses from patients or the environment. Dey and coworkers developed synthetic polymer nanogels that were able to bind herpes simplex virus 1 and block its ability to infect mammalian cells *in vitro*[71]. More recently, Pu and coworkers have fused an influenza-binding peptide to curli fibers and demonstrated their ability to remove virus particles from contaminated water[78]. Strategies based on feeding of VHHs or VHH multimers have also been reported and have shown promise for treating the pathology of enteric pathogen infections[18,19,22].

In our VHH generation efforts, initial attempts at identifying intimin-binding antibodies by panning on purified protein proved unsuccessful, an issue that was rectified by panning instead on intimin-expressing bacteria. While using a properly folded protein is an important consideration for the selection of any binding interaction, it is especially relevant for VHHs, for which binding is known to be particularly dependent on protein

conformation[48]. Our results strongly suggest that recombinant intimin attached to a plastic well does not accurately reflect the conformational state of bacterial surface intimin. Interestingly, plastic beads, when coated with a similar preparation of recombinant intimin, trigger actin pedestal formation when added to mammalian cells that harbor plasma membrane Tir[79], indicating that recombinant intimin can display biologic function under some circumstances.

3.5. Materials and Methods

3.5.1. Cell strains and plasmids

All strains and plasmids used in this study are summarized in Supplementary Table C1.

3.5.2. Bacterial culture

All *E. coli* and *C. rodentium* strains were cultured in LB broth at 37°C at 225 RPM, unless otherwise specified. EcN (PBP8) strains were streaked from frozen stock onto selective lysogeny broth (LB) agar plates and grown overnight at 37°C. Cultures were subsequently started from single colonies into 5 mL LB supplemented with 50 µg/mL kanamycin and grown overnight at 37°C with shaking at 225 RPM. The following day, overnight cultures were diluted 1:100 into 10 mL fresh media and grown at 37°C and 225 RPM, and protein expression was induced by adding 100 µM isopropyl ß-D-1-thiogalactopyranoside (IPTG). Induced cultures were grown overnight.

Shigella flexneri was streaked from frozen stock onto tryptic soy agar (TSA) plates supplemented with 0.02% Congo Red (CR) and grown overnight at 37°C. The following day, 3 colonies were used to inoculate 50 mL of tryptic soy broth (TSB) in a baffled flask. Only

colonies stained red by CR were used. The culture was grown at 37°C and 225 RPM to OD at 600 nm of 0.8-1.0, placed on ice upon reaching the desired OD, pelleted at 3500 RPM for 10 minutes at 4°C, and resuspended in 5 mL to obtain a 10x suspension.

3.5.3. Mammalian cell culture

HeLa cells (ATCC CCL-2) were maintained in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum (FBS) and 1% Penicillin-Streptomycin antibiotics in a 5% CO2 incubator at 37°C. For infection, 30,000 cells were seeded on 24-well plates in a volume of 0.5 mL/well. The next day, cells were gently washed with PBS before inoculating with bacteria. Vero cells (ATCC CCL-81) were grown in Eagle's Minimum Essential Medium (EMEM) supplemented with 10% FBS in a 5% CO2 incubator at 37°C. For Stx2 toxicity assays, Vero cells were seeded on 96-well plates one day prior to incubation with toxin.

3.5.4. Pathogen propagation

C. parvum oocysts, MD isolate originating from deer and passaged repeatedly in sheep and mice[80], were generated at Tufts University by propagation in CD-1 mice as described elsewhere[81], in compliance with study protocol No. G2017-107 approved by the Tufts University Institutional Animal Care Use Committee. Prior to excystation, oocysts were bleached on ice for 7 minutes using 5% dilution of commercial bleach (Clorox Original, The Clorox Company, CA). To remove bleach, oocysts were washed three times by suspension in PBS and centrifugation (18,000 × g, 2 min).

3.5.5. Purification of flagella

Flagella from REPEC (E22), EPEC (E2348/69), and E10 (O119:H6) were isolated as described previously[82], with slight modification. Briefly, a single colony was transferred into

5 mL LB broth and incubated overnight at 37°C with continuous shaking. The next day, the culture was diluted 1:100 into LB broth and grown at 37°C to OD_{600} of 0.5. 100 µL of the culture was plated onto the surface of eighty 100mm diameter LB agar plates and incubated for 24h at 37°C. 500 µL of PBS was then added to each plate and a glass slide was used to gently scrape bacteria from the agar plate. Bacteria were collected in a centrifuge bottle. To shear flagella from the bacteria, the centrifuge bottle was manually shaken for 2 min and then shaken for 5 mins at 4°C at 220 RPM. The bottle was then centrifuged at 7025 × g for 20 min at 4°C to remove cell debris. Bacteria-free supernatant was transferred to a new centrifuge bottle, which was further centrifuged at 25,402 × g for 1 hour at 4°C to precipitate flagella. To recover flagella, the supernatant was removed, and the pellet was resuspended in 500 µL of ice-cold PBS. To confirm that the purified flagella encompassed flagellin monomers of 60 kDa, flagella were visualized by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and by Western blotting using Rabbit anti-H6 flagella antibody. Note: EPEC E2348/69 produces fewer flagella filaments when grown in LB media[51]. Therefore, to maximize shearing of flagella from E2348/69, bacteria were either passed through a syringe and a 22-gauge needle or heat treated at 65°C for 30 mins.

3.5.6. Alpaca immunizations

Immunizations were performed essentially as described by Vrentas *et al.* [83]. Two different pairs of alpacas were each immunized in two separate rounds of immunization with various combinations of purified REPEC or EPEC flagella, and/or recombinant proteins MBP/EHEC intimin, or 6xHis/EHEC Tir. For each round of immunization, five successive multi-site subcutaneous injections were employed at about 3-week intervals. Blood was obtained for lymphocyte preparation 3-5 days after the fifth immunization and RNA was

prepared from lymphocytes using the RNeasy kit (Qiagen, Valencia, CA). A VHH-display phage library was prepared essentially as described previously[70] following each of the rounds of alpaca immunization, yielding libraries with complexities of about 1-2 x 10^7 independent clones, and >95% containing VHH inserts.

3.5.7. Identification and purification of VHHs

Phage library panning methods have been previously described[84]. Typically, the virulence factor proteins were coated onto plastic at 10 μg/mL of target in the first panning round, followed by a second round of panning at high stringency, with virulence factor proteins coated at 1 μg/mL, and using a 10-fold lower titer of input phage, shorter binding times, and longer washes. In some cases, the virulence factor targets were captured by previously coating the selection vessels with VHHs recognizing the target or its fusion partner. Random clones from the selected populations were then screened by ELISA for expression of VHHs that bound to the virulence factor targets. Clones producing the strongest signals or showing broader target specificity were characterized by DNA fingerprinting. The sequences of the DNAs encoding the VHHs displaying the strongest ELISA signals were obtained, and based on sequence homology, one VHH representing unique homology groups (having no evidence of a common B cell clonal origin) was selected for expression and characterization. These VHHs were expressed individually in pET32 vectors and purified as recombinant *E. coli* thioredoxin fusions with a carboxy-terminal E-tag, as previously described[84].

3.5.8. Dilution ELISA

ELISAs were performed using Nunc Maxisorp 96 well plates (Thermo Fisher Scientific). Virulence factor targets were typically coated overnight at 4°C, 1 μg/mL in PBS, then

blocked for at least an hour at 37°C with 4% milk in PBS, 0.1% Tween. For capture ELISAs, plates were first coated with 5 µg/mL VHHs that recognized the virulence factor or its fusion partner. The captured VHHs lacked both the thioredoxin partner and E-tag. After blocking, the virulence factor was then incubated at 1 µg/mL for one hour at 37°C with 4% milk in PBS, 0.1% Tween and washed. Dilution ELISAs were then initiated by diluting the VHH (expressed in vector pET-32 with an amino terminal thioredoxin and a carboxyl terminal E-tag) to 125 nM and performing serial dilutions of 1:5. After incubation for one hour at 37°C, plates were washed and then incubated with 1:10,000 rabbit HRP/anti-E-tag (Bethyl Laboratories) for one hour, washed, developed with TMB (Sigma Aldrich) as recommended by the manufacturer and measured for absorbance at 450 nm.

3.5.9. ELISA measuring the effect of anti-Tir VHHs on the intimin-Tir interaction

The ability of anti-Tir VHHs derived from EHEC to interfere with intimin-Tir binding was measured by ELISA. High-binding assay plates (Corning) were coated with 5 µg/mL of recombinant his-tagged Tir diluted in 1x coating buffer (50 mM Na2CO3, 50 mM NaHCO3, pH 9.6) in a volume of 100 µL per well and incubated overnight at 4°C. Plates were washed three times with 300 µL of wash buffer (0.05% Tween in PBS) and then blocked with BSA (3% in PBS) for 2 hours at room temperature (RT). Plates were washed and 100 µL of 500 nM anti-Tir VHH were added to each well. 0.1% BSA was used as a negative control. Plates were incubated at RT for 2 hours or at 4°C overnight. Wells were then probed with 150 nM GST-tagged intimin or with GST alone, and incubated at RT for 2 hours or at 4°C overnight. Plates were washed again and then fixed with 3.7% paraformaldehyde at RT for 20 mins at 4°C. Following another wash, plates were blocked with 5% milk in PBS for 30 min at RT. After washing, plates were incubated with goat anti-GST (GE Healthcare) for an hour and GST binding was detected kinetically using an alkaline-phosphatase-linked rabbit

anti-goat IgG secondary antibody (diluted 1:2000 in 0.1% BSA/PBS). Binding of the secondary antibody was detected colorimetrically (AP substrate N1891, Sigma Aldrich) at 405 nm, and the average reaction rate (V_{mean}) was calculated.

3.5.10. REPEC motility assay

Motility assays were performed as described previously[70], with slight modification, to measure the ability of anti-Fla VHHs to inhibit Rabbit Enteropathogenic Escherichia coli (REPEC) motility. Briefly, REPEC cultures were streaked on LB agar plates and incubated for 16 hours at 37°C. The next day, a single colony was transferred into 5 mL LB broth and incubated overnight at 37°C with continuous shaking. On the following day, the culture was diluted 1:50 into LB broth and grown at 37°C with continuous shaking to an OD_{600} of 0.5. A 1:1 dilution of bacteria and VHHs was then prepared (6 µL of bacterial culture was mixed with 6 µL of VHH concentrations ranging from 0 to 6.4 µM), mixed gently with a pipette, and incubated at 4°C for 2 h. The 12 µL mixture was then transferred to the center surface of a 0.3% semi solid agar plate and incubated for 24 h at room temperature. The diameter of bacterial growth was measured by first placing the plate on a dark background to enhance the contrast between bacterial growth and the agar medium. A metric scale ruler was then used to measure the growth diameter. Images were captured using a Syngene imager.

To test the ability of EcN-produced CsgA-VHH to inhibit REPEC motility, the above procedure was modified to embed induced EcN suspensions in the semi-solid agar medium. For each EcN variant, 50 mL of culture were grown and induced as previously described. Each culture was pelleted and resuspended in 1% tryptone 0.5% NaCl media supplemented with 200 µg/mL carbenicillin to prevent EcN growth. Each plate was filled with 12.5 mL EcN suspension followed by 12.5 mL 0.6% agar media, for a final plate composition of 1%

tryptone, 0.5% NaCl, 0.3% agar, 100 µg/mL carbenicillin and CgsA-VHH dispersed throughout the volume of the plate. For this assay, REPEC was transformed with the KAG plasmid to confer carbenicillin resistance as well as constitutive GFP expression to better distinguish it from EcN biomass. Images were acquired at several time points over a 48-hour period using a GelDoc imaging system (Bio-Rad).

3.5.11. EPEC pedestal assay

The ability of anti-intimin and anti-Tir VHHs to inhibit EPEC pedestal formation was assessed after infection of HeLa cells, as described previously[85], with slight modification. Briefly, 30,000 HeLa cells were inoculated into the wells of 24-well plates (Invitro Scientific,) and incubated overnight at 37°C in an incubator with 5% CO2. On the same day, a single EPEC colony was inoculated into 5 mL DMEM in 100mM HEPES medium (pH 7.4) and incubated overnight in 5% CO2 at 37°C without shaking. The next day, the EPEC culture was diluted 1:16 into new infection medium (0.6 µL EPEC added to 9.4 µL media containing DMEM, 20 mM HEPES, and 3.5% FBS; pH 7.4), and 3.33 µL of the EPEC suspension were incubated either alone or with 100 nM anti-Tir or anti-intimin VHH in 0.5 mL DMEM at 4°C for 2 hours, in 1.5 mL Eppendorf tubes on a rocker. HeLa cell monolayers were then washed with 0.5 mL PBS, and EPEC suspensions were added to the monolayers. Plates were then centrifuged at 500 RPM for 5 min and incubated for 3 h at 37°C in a 5% CO2 incubator. Next, cells were washed twice with PBS, fixed with 0.5 mL 2.5% paraformaldehyde in PBS for 10 mins at RT on a shaker, washed twice with PBS for 5 min on a shaker at RT, permeabilized with 0.5 mL of 0.1% TritonX-100 for 5 min, and washed twice again before staining with DAPI (Thermo Fisher Scientific) and Alexa Fluor-488 Phalloidin (Thermo Fisher Scientific) at RT for 1.5 hours. Monolayers were then

washed and 7 µL prolong gold anti-fade reagent (Thermo Fisher Scientific) was used to mount coverslips on wells before imaging with a fluorescent microscope.

EPEC pedestal formation was blindly scored, as follows; 1: Very few pedestals are present on the edges of the wells, 2: More pedestals present, only at the edges of the well, 3: Most cells have no pedestals, but a few pedestals present in the center and edges of wells, 4: Most cells have pedestals, but a few empty cells are present, 5: The majority of the cells have pedestals. Using the above numbering criteria, pedestals were scored blindly by a second researcher.

3.5.12. CsgA-VHH plasmid construction and cloning

The cloning of the synthetic curli operon *csgBACEFG* onto the pL6FO vector was described in detail elsewhere[86]. DNA sequences of desired VHHs and corresponding primers were synthesized by and purchased from Integrated DNA Technologies. Plasmid construction was carried out using Gibson Assembly[87].

3.5.13. Quantitative Congo Red binding assay

Curli fiber formation was quantified using a Congo Red binding assay as previously described[44]. Briefly, 1 mL of induced EcN CsgA-VHH culture was pelleted at 4000 × g for 10 minutes at room temperature and resuspended in a 25 µM Congo Red PBS solution. After a 10-minute incubation, the cell suspension was pelleted again, and the unbound Congo Red dye was quantified by measuring the supernatant absorbance at 490 nm. The signal was subtracted from a Congo Red blank, divided by the culture's OD_{600} measurement to reflect curli production per cell, and normalized with respect to a EcN CsgA (no VHH) positive control.

3.5.14. Electron microscopy

Field emission scanning electron microscope (FESEM) samples were prepared by fixing with 2% (w/v) glutaraldehyde and 2% (w/v) paraformaldehyde at room temperature, overnight. The samples were gently washed with water, and the solvent was gradually exchanged with ethanol with an increasing ethanol 15-minute incubation step gradient (25, 50, 75 and 100% (v/v) ethanol). The samples were dried in a critical point dryer, placed onto SEM sample holders using silver adhesive (Electron Microscopy Sciences) and sputtered until they were coated in a 10-20 nm layer of Pt/Pd. Images were acquired using a Zeiss Ultra55 FESEM equipped with a field emission gun operating at 5-10 kV.

3.5.15. GFP pull-down assay

For each condition, 1 mL of induced overnight culture was centrifuged at 4000 × g for 10 minutes. The supernatant was aspirated, and the pellets were resuspended in 150 nM GFP in fasted-state simulated colonic fluid, which was prepared as described by Vertzoni *et al.* [88]. The cells were incubated on a shaking platform (225 RPM) at 37°C for 15 minutes and pelleted again. The GFP remaining in solution was assayed by measuring the fluorescent signal (485nm/528nm) using a plate reader (Spectramax M5, Molecular Devices). GFP concentration was estimated based on a calibration curve using known GFP concentrations.

3.5.16. Shiga toxin pull-down assay

Induced EcN cultures were pelleted at 4000 × g for 10 minutes and resuspended in 10 ng/mL of Stx2 in PBS. The bacterial suspensions were then serially diluted tenfold (from 1:10 to $1:10^4$) in 10 ng/mL Stx2 PBS solution, maintaining a constant Stx2 concentration. The suspensions were incubated at 37°C on a 225 RPM rotating platform for 1 hour and pelleted again. For each condition, 10 µL of supernatant were added to 90 µL of Vero cell

medium in its corresponding well. After a 48-hour incubation at 37°C with 5% CO_2, 10 µL of PrestoBlue Cell Viability Reagent (Thermo Fisher Scientific) was added into each well, followed by a 10-minute incubation and measurement of fluorescent signal at 560nm/590nm.

3.5.17. REPEC and REHEC aggregation assays

REPEC, REHEC and EcN were cultured as previously described. Cultures were pelleted and resuspended in PBS to obtain 3x, 1x, 0.3x or 0.1x suspensions as compared to the original culture density. Cell suspensions were subsequently mixed and added into 96-well conical-bottom microwell plates (Thermo Fisher Scientific) and allowed to settle overnight at room temperature prior to imaging.

3.5.18. Generation of anti-IpaD VHH trimer

VHH heterotrimer was designed and generated as previously described[42,89]. Briefly, DNA encoding the 20ipaD, JMJ-F5, and JPS-G3 VHHs[20] was joined in frame downstream of thioredoxin (trx). VHHs sequences were separated by 15-amino acid flexible glycine-serine linkers ($(GGGGS)_3$), and a C-terminal E-tag epitope was included downstream of VHH sequences.

3.5.19. *Shigella* contact-mediated hemolysis assay

To determine the ability of CsgA-VHH to inhibit *Shigella* virulence activity, a contact-mediated hemolysis assay was carried out as previously described[62], with slight modification. Prior to exposure of *S. flexneri* to red blood cells, induced EcN cultures were pelleted and resuspended in PBS to obtain a 10x concentrated cell suspension, which was subsequently serially diluted to yield 5x, 2.5x, 1.25x and 0.625x suspensions. The pathogen was then incubated for 30 min at room temperature with either the EcN suspensions or the

abovementioned anti-IpaD VHH heterotrimer (trx/20ipaD/JMJ-F5/JPS-G3/E) as a positive control, in concentrations between 3.2-2000 nM.

3.5.20. Preparation of *Cryptosporidium* lysate

Pre-bleached *C. parvum* oocysts were excysted in 0.75% taurocholic acid suspension in PBS for 1h at 37°C. Following centrifugation (18,000 × g, 2 min), supernatant was collected and the pelleted sample consisting of sporozoites, unexcysted oocysts and oocyst shells was then sonicated (Qsonica CL5, Qsonica Sonicators, USA) with thirty cycles, 20 seconds each. Sonicated pellet was resuspended in supernatant and saved as '*C. parvum* whole lysate'. The concentration of the antigen fractions was determined by measurement of optical density using a Nanodrop instrument (ND-1000, NanoDrop Technologies).

3.5.21. Pull-down of *C. parvum* antigens using EcN

A variety of modified pull-down studies utilizing the principles of ELISA, Western blot and immunofluorescence were applied to test the ability of anti-gp900 VHHs fused to EcN curli to bind their targets. All experiments used a nonspecific control EcN construct which expressed curli in fusion with a VHH targeting the green fluorescent protein (CsgA-αGFP).

For the ELISA, the goal was to pull-down *C. parvum* antigens by EcN immobilized to a plastic surface. Briefly, 100 µl of induced overnight EcN cultures expressing CsgA-αgp900-1 and -2 were coated on 96-well MaxiSorb plates at 2x concentration and incubated overnight at 4°C. The following day, plates were washed with TBS-0.1% Tween and blocked with 4% milk-TBS-0.1% Tween solution for 1 h at 37°C. Plates were washed and *C. parvum* whole lysate was applied in 2-fold dilutions starting with 50 µg/mL concentration, and then incubated for 1 h at 37°C. After washing, specifically bound *C. parvum* antigen was incubated with a second E-tagged detection VHH that binds to the same *C. parvum* antigen

recognized by the EcN displayed VHH, but to a non-competing epitope, at 1 μg/mL for 1 h at 37°C. Plates were then washed and incubated with an anti-E-tag HRP antibody (Bethyl Laboratories) at 1:10,000 for 1 h at 37°C. Plates were washed a final time and OPD was added to each well for 20 minutes. The reaction was stopped with 1 M H2SO4 and absorbance was measured at 490 nm using a microplate reader.

For the Western blot, the goal was to quantify depletion of the target in the soluble whole lysate of *C. parvum* after incubation with EcN displaying an anti-gp900 VHH and removal from the solution by centrifugation. For target pull down, 50 μL of the induced (for VHH display) and blocked EcN was suspended in PBS at 2x concentration and incubated with 30 μg of *C. parvum* whole lysate for 1 h with rotation at room temperature. Samples were centrifuged (5000 ×g, 1 min) and supernatant was collected for analysis. Fifteen μL supernatant aliquots of supernatant were diluted with 4xLDS buffer (Novagen) to achieve 1x concentration and denatured at 70°C for 10 minutes. Samples were loaded into the wells of 4-12% Bis-Tris gel (Novex) and electrophoresed in 1x MOPS buffer at 100 V for 10 minutes and then at 200 V for 40 minutes. The gel was transferred on the nitrocellulose membrane using a wet transfer system (395 mA, 4h). Membranes were blocked with 4% milk-TBS 0.1% Tween for 1h and washed with TBS-T before blotting with a second detection VHH (recognizing a non-competing epitope on the target) at 1 μg/ml for 1 h with rotation. Membranes were then washed and incubated with a secondary anti-E-tag HRP antibody at 1:5,000 dilutions for 1 h with rotation. Western blots were developed using chemiluminescent substrate (GE Healthcare) and imaged using a ChemiDoc system (Bio-Rad). A densitometry analysis was performed using Image Lab software to report the percent of target band depletion as normalized to the loading control.

3.5.22. Immunofluorescence imaging of EcN attachment to *C. parvum*

Immunofluorescent imaging was used to quantify EcN bacteria attached to *C. parvum* parasite and its trails immobilized on the surface. Pre-bleached *C. parvum* oocysts were suspended in 0.75% taurocholic acid and excysted in a 37°C water bath for 30 minutes to release sporozoites. Aliquots of excysted 10,000 oocysts were transferred onto poly-L-lysine slides (Chromaview) and incubated for another 30 minutes at 37°C under humidified conditions to allow for further excystation and gliding of sporozoites. Slides were then dried, fixed with 4% paraformaldehyde at room temperature (20 min) and washed with PBS. Such prepared parasites were then probed with 200 µL of 2x EcN suspensions and incubated for 1h at room temperature, after which they were washed with PBS to remove unbound EcN. To detect EcN bound to the sporozoites and trails, slides were probed with anti-LPS Mab (ThermoFisher Scientific) at 1:200 dilution, followed by an anti-mouse IgG Alexa Fluor 568 antibody (Invitrogen) at 1:500 dilution, both incubated for 1h at room temperature. Sporozoites were counterstained with an E-tagged VHH targeting gp900 at the apical complex and trails (CsgA-αgp900-2) at 1 µg/mL concentration, followed by an anti-E-tag-FITC antibody (Bethyl Laboratories) at 1:100 dilution, both incubated for 1h at room temperature. Lastly, slides were washed, dried, and mounted with antifade medium. Fluorescing sporozoites were imaged under epifluorescence (Nikon Eclipse Ti-E microscope, Nikon Instruments Inc.). The number of fluorescent foci was quantified using ImageJ 1.48v particle analyzer (U.S. National Institutes of Health, Bethesda, Maryland, USA).

3.5.23. Statistical analysis

All statistical analyses were performed using Prism 9.1.0 (GraphPad Software). Data are presented as mean ± standard error of mean (SEM), unless otherwise specified. Statistical significance was assessed using one-way or two-way analysis of variance (ANOVA), followed by Welsh's t-test, as described in figure legends.

3.6. Acknowledgements

This work was done in collaboration with Dr. Yousuf Aqeel, Dr. Jacqueline M. Tremblay, Dr. Justyna J. Jaskiewicz, Anishma Shrestha, James N. Lee, Shenglan Hu, Dr. Xi Qian, Loranne Magoun, Prof. Abhineet Sheoran, Prof. Daniela Bedenice, Colter Giem, Dr. Avinash Manjula-Basavanna, Dr. Marcia S. Osburne, Prof. Saul Tzipori, Prof. Charles B. Shoemaker, Prof. John M. Leong, and Prof. Neel S. Joshi. This work was supported by the National Institutes of Health (1R01DK110770-01A1, R01DK113599-01) and the Bill and Melinda Gates Foundation (OPP1172434) (CBS/JML).

3.7. References

1. Kotloff, K. L., Blackwelder, W. C., Nasrin, D., Nataro, J. P., Farag, T. H., van Eijk, A., ... & Levine, M. M. (2012). The Global Enteric Multicenter Study (GEMS) of diarrheal disease in infants and young children in developing countries: epidemiologic and clinical methods of the case/control study. *Clinical infectious diseases*, *55*(suppl_4), S232-S245.

2. Kotloff, K. L., Nataro, J. P., Blackwelder, W. C., Nasrin, D., Farag, T. H., Panchalingam, S., ... & Levine, M. M. (2013). Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *The Lancet*, *382*(9888), 209-222.

3. Nicolini, G., Sperotto, F., & Esposito, S. (2014). Combating the rise of antibiotic resistance in children. *Minerva Pediatr*, *66*(1), 31-9.

4. Walker, C. L. F., Perin, J., Aryee, M. J., Boschi-Pinto, C., & Black, R. E. (2012). Diarrhea incidence in low-and middle-income countries in 1990 and 2010: a systematic review. *BMC public health*, *12*(1), 1-7.

5. Kotloff, K. L., Nasrin, D., Blackwelder, W. C., Wu, Y., Farag, T., Panchalingham, S., ... & Levine, M. M. (2019). The incidence, aetiology, and adverse clinical consequences of less severe diarrhoeal episodes among infants and children residing in low-income and middle-income countries: a 12-month case-control study as a follow-on to the Global Enteric Multicenter Study (GEMS). *The Lancet Global Health*, *7*(5), e568-e584.

6. Levine, M. M., Nasrin, D., Acácio, S., Bassat, Q., Powell, H., Tennant, S. M., ... & Kotloff, K. L. (2020). Diarrhoeal disease and subsequent risk of death in infants and children residing in low-income and middle-income countries: analysis of the GEMS case-control study and 12-month GEMS-1A follow-on study. *The Lancet Global Health*, *8*(2), e204-e214.

7. Moon, H. W., Whipp, S. C., Argenzio, R. A., Levine, M. M., & Giannella, R. A. (1983). Attaching and effacing activities of rabbit and human enteropathogenic Escherichia coli in pig and rabbit intestines. *Infection and immunity*, *41*(3), 1340-1351.

8. Kaper, J. B., Nataro, J. P., & Mobley, H. L. (2004). Pathogenic escherichia coli. *Nature reviews microbiology*, *2*(2), 123-140.

9. Lai, Y., Rosenshine, I., Leong, J. M., & Frankel, G. (2013). Intimate host attachment: enteropathogenic and enterohaemorrhagic E scherichia coli. *Cellular microbiology*, *15*(11), 1796-1808.

10. Mallick, E. M., Garber, J. J., Vanguri, V. K., Balasubramanian, S., Blood, T., Clark, S., ... & Leong, J. M. (2014). The ability of an attaching and effacing pathogen to trigger localized actin assembly contributes to virulence by promoting mucosal attachment. *Cellular microbiology*, *16*(9), 1405-1424.

11. Gyles, C. L. (2007). Shiga toxin-producing Escherichia coli: an overview. *Journal of animal science*, *85*(suppl_13), E45-E62.

12. Marder, E. P., Griffin, P. M., Cieslak, P. R., Dunn, J., Hurd, S., Jervis, R., ... & Geissler, A. L. (2018). Preliminary incidence and trends of infections with pathogens transmitted commonly through food—foodborne diseases active surveillance network, 10 US Sites, 2006–2017. *Morbidity and Mortality Weekly Report*, *67*(11), 324.

13. Fatima, R., & Aziz, M. (2020). Enterohemorrhagic Escherichia Coli (EHEC). *StatPearls [Internet]*.

14. Schauer, D. B., & Falkow, S. (1993). Attaching and effacing locus of a Citrobacter freundii biotype that causes transmissible murine colonic hyperplasia. *Infection and immunity*, *61*(6), 2486-2492.

15. Mundy, R., MacDonald, T. T., Dougan, G., Frankel, G., & Wiles, S. (2005). Citrobacter rodentium of mice and man. *Cellular microbiology*, *7*(12), 1697-1706.

16. Gaytán, M. O., Martínez-Santos, V. I., Soto, E., & González-Pedrajo, B. (2016). Type three secretion system in attaching and effacing pathogens. *Frontiers in cellular and infection microbiology*, *6*, 129.

17. Maffey, L., Vega, C. G., Parreno, V., & Garaicoechea, L. (2015). Controlling Rotavirus-associated diarrhea: Could single-domain antibody fragments make the difference?. *Revista Argentina de microbiologia*, *47*(4), 368-379.

18. Virdi, V., Palaci, J., Laukens, B., Ryckaert, S., Cox, E., Vanderbeke, E., ... & Callewaert, N. (2019). Yeast-secreted, dried and food-admixed monomeric IgA prevents gastrointestinal infection in a piglet model. *Nature biotechnology*, *37*(5), 527-530.

19. Sheoran, A. S., Dmitriev, I. P., Kashentseva, E. A., Cohen, O., Mukherjee, J., Debatis, M., ... & Tzipori, S. (2015). Adenovirus vector expressing Stx1/Stx2-neutralizing agent protects piglets infected with Escherichia coli O157: H7 against fatal systemic intoxication. *Infection and immunity*, *83*(1), 286-291.

20. Barta, M. L., Shearer, J. P., Arizmendi, O., Tremblay, J. M., Mehzabeen, N., Zheng, Q., ... & Picking, W. L. (2017). Single-domain antibodies pinpoint potential targets within Shigella invasion plasmid antigen D of the needle tip complex for inhibition of type III secretion. *Journal of Biological Chemistry*, *292*(40), 16677-16687.

21. Ruano-Gallego, D., Yara, D. A., Di Ianni, L., Frankel, G., Schüller, S., & Fernández, L. Á. (2019). A nanobody targeting the translocated intimin receptor inhibits the attachment of enterohemorrhagic E. coli to human colonic mucosa. *PLoS pathogens*, *15*(8), e1008031.

22. Schmidt, D. J., Beamer, G., Tremblay, J. M., Steele, J. A., Kim, H. B., Wang, Y., ... & Tzipori, S. (2016). A tetraspecific VHH-based neutralizing antibody modifies disease outcome in three animal models of Clostridium difficile infection. *Clinical and Vaccine Immunology*, *23*(9), 774-784.

23. Vanmarsenille, C., Del Olmo, I. D., Elseviers, J., Ghassabeh, G. H., Moonens, K., Vertommen, D., ... & De Greve, H. (2017). Nanobodies targeting conserved epitopes on the major outer membrane protein of Campylobacter as potential tools for control of Campylobacter colonization. *Veterinary research*, *48*(1), 1-14.

24. Riazi, A., Strong, P. C., Coleman, R., Chen, W., Hirama, T., van Faassen, H., ... & Ghahroudi, M. A. (2013). Pentavalent single-domain antibodies reduce Campylobacter jejuni motility and colonization in chickens. *PloS one*, *8*(12), e83928.

25. Tremblay, J. M., Mukherjee, J., Leysath, C. E., Debatis, M., Ofori, K., Baldwin, K., ... & Shoemaker, C. B. (2013). A single VHH-based toxin-neutralizing agent and an effector antibody protect mice against challenge with Shiga toxins 1 and 2. *Infection and immunity*, *81*(12), 4592-4603.

26. Vega, C. G., Bok, M., Vlasova, A. N., Chattha, K. S., Gómez-Sebastián, S., Nuñez, C., ... & Parreño, V. (2013). Recombinant monovalent llama-derived antibody fragments (VHH) to rotavirus VP6 protect neonatal gnotobiotic piglets against human rotavirus-induced diarrhea. *PLoS Pathog*, *9*(5), e1003334.

27. Maffey, L., Vega, C. G., Miño, S., Garaicoechea, L., & Parreño, V. (2016). Anti-VP6 VHH: an experimental treatment for rotavirus A-associated disease. *PLoS One*, *11*(9), e0162351.

28. Amcheslavsky, A., Wallace, A. L., Ejemel, M., Li, Q., McMahon, C. T., Stoppato, M., ... & Wang, Y. (2021). Anti-CfaE nanobodies provide broad cross-protection against major pathogenic enterotoxigenic Escherichia coli strains, with implications for vaccine design. *Scientific reports*, *11*(1), 1-15.

29. Muyldermans, S. (2020). Applications of Nanobodies. *Annual Review of Animal Biosciences*, *9*.

30. Muyldermans, S. (2013). Nanobodies: natural single-domain antibodies. *Annual review of biochemistry*, *82*, 775-797.

31. Gurbatri, C. R., Lia, I., Vincent, R., Coker, C., Castro, S., Treuting, P. M., ... & Danino, T. (2020). Engineered probiotics for local tumor delivery of checkpoint blockade nanobodies. *Science translational medicine*, *12*(530).

32. Chiabai, M. J., Almeida, J. F., de Azevedo, M. G. D., Fernandes, S. S., Pereira, V. B., de Castro, R. J. A., ... & Brigido, M. M. (2019). Mucosal delivery of Lactococcus lactis carrying an anti-TNF scFv expression vector ameliorates experimental colitis in mice. *BMC biotechnology*, *19*(1), 1-12.

33. Afchangi, A., Latifi, T., Jalilvand, S., Marashi, S. M., & Shoja, Z. (2021). Combined use of lactic-acid-producing bacteria as probiotics and rotavirus vaccine candidates expressing virus-specific proteins. *Archives of Virology*, 1-12.

34. Ou, B., Yang, Y., Tham, W. L., Chen, L., Guo, J., & Zhu, G. (2016). Genetic engineering of probiotic Escherichia coli Nissle 1917 for clinical application. *Applied microbiology and biotechnology*, *100*(20), 8693-8699.

35. Kurtz, C. B., Millet, Y. A., Puurunen, M. K., Perreault, M., Charbonneau, M. R., Isabella, V. M., ... & Miller, P. F. (2019). An engineered E. coli Nissle improves hyperammonemia and survival in mice and shows dose-dependent exposure in healthy humans. *Science translational medicine*, *11*(475).

36. Kruis, W., Frič, P., Pokrotnieks, J., Lukáš, M., Fixa, B., Kaščák, M., ... & Schulze, J. (2004). Maintaining remission of ulcerative colitis with the probiotic Escherichia coli Nissle 1917 is as effective as with standard mesalazine. *Gut*, *53*(11), 1617-1623.

37. Altenhoefer, A., Oswald, S., Sonnenborn, U., Enders, C., Schulze, J., Hacker, J., & Oelschlaeger, T. A. (2004). The probiotic Escherichia coli strain Nissle 1917 interferes with invasion of human intestinal epithelial cells by different enteroinvasive bacterial pathogens. *FEMS Immunology & Medical Microbiology*, *40*(3), 223-229.

38. Hancock, V., Dahl, M., & Klemm, P. (2010). Probiotic Escherichia coli strain Nissle 1917 outcompetes intestinalpathogens during biofilm formation. *Journal of medical microbiology*, *59*(4), 392-399.

39. Blum-Oehler, G., Oswald, S., Eiteljörge, K., Sonnenborn, U., Schulze, J., Kruis, W., & Hacker, J. (2003). Development of strain-specific PCR reactions for the detection of the probiotic Escherichia coli strain Nissle 1917 in fecal samples. *Research in microbiology*, *154*(1), 59-66.

40. Kurtz, C., Denney, W. S., Blankstein, L., Guilmain, S. E., Machinani, S., Kotula, J., ... & Brennan, A. M. (2018). Translational Development of Microbiome-Based Therapeutics: Kinetics of E. coli Nissle and Engineered Strains in Humans and Nonhuman Primates. *Clinical and translational science*, *11*(2), 200-207.

41. Gibbs, W. W. (2005). Nanobodies. *Scientific American*, *293*(2), 78-83.

42. Mukherjee, J., Tremblay, J. M., Leysath, C. E., Ofori, K., Baldwin, K., Feng, X., ... & Shoemaker, C. B. (2012). A novel strategy for development of recombinant antitoxin therapeutics tested in a mouse botulism model. *PloS one*, *7*(1), e29941.

43. Piñero-Lambea, C., Bodelón, G., Fernández-Periáñez, R., Cuesta, A. M., Álvarez-Vallina, L., & Fernández, L. A. (2015). Programming controlled adhesion of E. coli to target surfaces, cells, and tumors with synthetic adhesins. *ACS synthetic biology*, *4*(4), 463-473.

44. Praveschotinunt, P., Duraj-Thatte, A. M., Gelfat, I., Bahl, F., Chou, D. B., & Joshi, N. S. (2019). Engineered E. coli Nissle 1917 for the delivery of matrix-tethered therapeutic domains to the gut. *Nature communications*, *10*(1), 1-14.

45. Nguyen, P. Q., Botyanszki, Z., Tay, P. K. R., & Joshi, N. S. (2014). Programmable biofilm-based materials from engineered curli nanofibres. *Nature communications*, *5*(1), 1-10.

46. Connor, E. F., Lees, I., & Maclean, D. (2017). Polymers as drugs—Advances in therapeutic applications of polymer binding agents. *Journal of Polymer Science Part A: Polymer Chemistry*, *55*(18), 3146-3157.

47. Liu, S., Maheshwari, R., & Kiick, K. L. (2009). Polymer-based therapeutics. *Macromolecules*, *42*(1), 3-13.

48. Pardon, E., Laeremans, T., Triest, S., Rasmussen, S. G., Wohlkönig, A., Ruf, A., ... & Steyaert, J. (2014). A general protocol for the generation of Nanobodies for structural biology. *Nature protocols*, *9*(3), 674-693.

49. Luo, Y., Frey, E. A., Pfuetzner, R. A., Creagh, A. L., Knoechel, D. G., Haynes, C. A., ... & Strynadka, N. C. (2000). Crystal structure of enteropathogenic Escherichia coli intimin–receptor complex. *Nature*, *405*(6790), 1073-1077.

50. Weikum, J., Kulakova, A., Tesei, G., Yoshimoto, S., Jægerum, L. V., Schütz, M., ... & Morth, J. P. (2020). The extracellular juncture domains in the intimin passenger adopt a constitutively extended conformation inducing restraints to its sphere of action. *Scientific reports*, *10*(1), 1-18.

51. Girón, J. A., Torres, A. G., Freer, E., & Kaper, J. B. (2002). The flagella of enteropathogenic Escherichia coli mediate adherence to epithelial cells. *Molecular microbiology*, *44*(2), 361-379.

52. Praveschotinunt, P., Dorval Courchesne, N. M., den Hartog, I., Lu, C., Kim, J. J., Nguyen, P. Q., & Joshi, N. S. (2018). Tracking of engineered bacteria in vivo using nonstandard amino acid incorporation. *ACS synthetic biology*, *7*(6), 1640-1650.

53. Zhou, Y., Smith, D. R., Hufnagel, D. A., & Chapman, M. R. (2013). Experimental manipulation of the microbial functional amyloid called curli. In *Bacterial cell surfaces* (pp. 53-75). Humana Press, Totowa, NJ.

54. Rothbauer, U., Zolghadr, K., Tillib, S., Nowak, D., Schermelleh, L., Gahl, A., ... & Leonhardt, H. (2006). Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nature methods*, *3*(11), 887-889.

55. Hussack, G., Arbabi-Ghahroudi, M., van Faassen, H., Songer, J. G., Ng, K. K. S., MacKenzie, R., & Tanha, J. (2011). Neutralization of Clostridium difficile toxin A with single-domain antibodies targeting the cell receptor binding domain. *Journal of Biological Chemistry*, *286*(11), 8961-8976.

56. Jaskiewicz, J. J., Tremblay, J. M., Tzipori, S., & Shoemaker, C. B. (2021). Identification and characterization of a new 34 kDa MORN motif-containing sporozoite surface-exposed protein, Cp-P34, unique to Cryptosporidium. *International Journal for Parasitology*.

57. Melton-Celsa, A. R. (2015). Shiga toxin (Stx) classification, structure, and function. *Enterohemorrhagic Escherichia coli and Other Shiga Toxin-Producing E. coli*, 37-53.

58. Namba, K., Yamashita, I., & Vonderviszt, F. (1989). Structure of the core and central channel of bacterial flagella. *Nature*, *342*(6250), 648-654.

59. Bajunaid, W., Haidar-Ahmad, N., Kottarampatel, A. H., Ourida Manigat, F., Silué, N., F Tchagang, C., ... & Campbell-Valois, F. X. (2020). The T3SS of Shigella: Expression, Structure, Function, and Role in Vacuole Escape. *Microorganisms*, *8*(12), 1933.

60. Espina, M., Olive, A. J., Kenjale, R., Moore, D. S., Ausar, S. F., Kaminski, R. W., ... & Picking, W. L. (2006). IpaD localizes to the tip of the type III secretion system needle of Shigella flexneri. *Infection and immunity*, *74*(8), 4391-4400.

61. Dickenson, N. E., Zhang, L., Epler, C. R., Adam, P. R., Picking, W. L., & Picking, W. D. (2011). Conformational changes in IpaD from Shigella flexneri upon binding bile

salts provide insight into the second step of type III secretion. *Biochemistry*, *50*(2), 172-180.

62. Picking, W. L., Nishioka, H., Hearn, P. D., Baxter, M. A., Harrington, A. T., Blocker, A., & Picking, W. D. (2005). IpaD of Shigella flexneri is independently required for regulation of Ipa protein secretion and efficient insertion of IpaB and IpaC into host membranes. *Infection and immunity*, *73*(3), 1432-1440.

63. Blocker, A., Gounon, P., Larquet, E., Niebuhr, K., Cabiaux, V., Parsot, C., & Sansonetti, P. (1999). The tripartite type III secreton of Shigella flexneri inserts IpaB and IpaC into host membranes. *The Journal of cell biology*, *147*(3), 683-693.

64. Centers for Disease Control and Prevention (CDC). Cryptosporidiosis Summary Report — National Notifiable Diseases Surveillance System, United States, 2018. Atlanta, Georgia: U.S. Department of Health and Human Services, CDC, 2019.

65. Kiros, H., Nibret, E., Munshea, A., Kerisew, B., & Adal, M. (2015). Prevalence of intestinal protozoan infections among individuals living with HIV/AIDS at Felegehiwot Referral Hospital, Bahir Dar, Ethiopia. *International journal of infectious diseases*, *35*, 80-86.

66. Love, M. S., & McNamara, C. W. (2021). Phenotypic screening techniques for Cryptosporidium drug discovery. *Expert Opinion on Drug Discovery*, *16*(1), 59-74.

67. Barnes, D. A., Bonnin, A., Huang, J. X., Gousset, L., Wu, J., Gut, J., ... & Petersen, C. (1998). A novel multi-domain mucin-like glycoprotein of Cryptosporidium parvum mediates invasion. *Molecular and biochemical parasitology*, *96*(1-2), 93-110.

68. Kitov, P. I., Sadowska, J. M., Mulvey, G., Armstrong, G. D., Ling, H., Pannu, N. S., ... & Bundle, D. R. (2000). Shiga-like toxins are neutralized by tailored multivalent carbohydrate ligands. *Nature*, *403*(6770), 669-672.

69. Mulvey, G. L., Marcato, P., Kitov, P. I., Sadowska, J., Bundle, D. R., & Armstrong, G. D. (2003). Assessment in mice of the therapeutic potential of tailored, multivalent Shiga toxin carbohydrate ligands. *The Journal of infectious diseases*, *187*(4), 640-649.

70. Moayeri, M., Leysath, C. E., Tremblay, J. M., Vrentas, C., Crown, D., Leppla, S. H., & Shoemaker, C. B. (2015). A heterodimer of a VHH (variable domains of camelid heavy chain-only) antibody that inhibits anthrax toxin cell binding linked to a VHH antibody that blocks oligomer formation is highly protective in an anthrax spore challenge model. *Journal of Biological Chemistry*, *290*(10), 6584-6595.

71. Dey, P., Bergmann, T., Cuellar-Camacho, J. L., Ehrmann, S., Chowdhury, M. S., Zhang, M., ... & Azab, W. (2018). Multivalent flexible nanogels exhibit broad-spectrum antiviral activity by blocking virus entry. *ACS nano*, *12*(7), 6429-6442.

72. Duraj-Thatte, A. M., Courchesne, N. M. D., Praveschotinunt, P., Rutledge, J., Lee, Y., Karp, J. M., & Joshi, N. S. (2019). Genetically programmable self-regenerating bacterial hydrogels. *Advanced Materials*, *31*(40), 1901826.

73. National Library of Medicine (U.S.). (2018, April 17 – 2019, June 21). *Safety and Tolerability of SYNB1618 in Healthy Adult Volunteers and Adult Subjects With*

Phenylketonuria (PKU). Identifier NCT03516487. https://clinicaltrials.gov/ct2/show/NCT03516487

74. Maltby, R., Leatham-Jensen, M. P., Gibson, T., Cohen, P. S., & Conway, T. (2013). Nutritional basis for colonization resistance by human commensal Escherichia coli strains HS and Nissle 1917 against E. coli O157: H7 in the mouse intestine. *PloS one*, *8*(1), e53957.

75. Jacobi, C. A., & Malfertheiner, P. (2011). Escherichia coli Nissle 1917 (Mutaflor): new insights into an old probiotic bacterium. *Digestive diseases*, *29*(6), 600-607.

76. Van Gerven, N., Klein, R. D., Hultgren, S. J., & Remaut, H. (2015). Bacterial amyloid formation: structural insights into curli biogensis. *Trends in microbiology*, *23*(11), 693-706.

77. Van Gerven, N., Goyal, P., Vandenbussche, G., De Kerpel, M., Jonckheere, W., De Greve, H., & Remaut, H. (2014). Secretion and functional display of fusion proteins through the curli biogenesis pathway. *Molecular microbiology*, *91*(5), 1022-1035.

78. Pu, J., Liu, Y., Zhang, J., An, B., Li, Y., Wang, X., ... & Zhong, C. (2020). Virus disinfection from environmental water sources using living engineered biofilm materials. *Advanced Science*, *7*(14), 1903558.

79. Liu, H., Magoun, L., Luperchio, S., Schauer, D. B., & Leong, J. M. (1999). The Tir-binding region of enterohaemorrhagic Escherichia coli intimin is sufficient to trigger actin condensation after bacterial-induced host cell signalling. *Molecular microbiology*, *34*(1), 67-81.

80. Okhuysen, P. C., Rich, S. M., Chappell, C. L., Grimes, K. A., Widmer, G., Feng, X., & Tzipori, S. (2002). Infectivity of a Cryptosporidium parvum isolate of cervine origin for healthy adults and interferon-γ knockout mice. *The Journal of infectious diseases*, *185*(9), 1320-1325.

81. Widmer, G., Tzipori, S., Fichtenbaum, C. J., & Griffiths, J. K. (1998). Genotypic and phenotypic characterization of Cryptosporidium parvum isolates from people with AIDS. *Journal of Infectious Diseases*, *178*(3), 834-840.

82. Tacket, C. O., Maneval, D. R., & Levine, M. M. (1987). Purification, morphology, and genetics of a new fimbrial putative colonization factor of enterotoxigenic Escherichia coli O159: H4. *Infection and immunity*, *55*(5), 1063-1069.

83. Vrentas, C. E., Moayeri, M., Keefer, A. B., Greaney, A. J., Tremblay, J., O'Mard, D., ... & Shoemaker, C. B. (2016). A diverse set of single-domain antibodies (VHHs) against the anthrax toxin lethal and edema factors provides a basis for construction of a bispecific agent that protects against anthrax infection. *Journal of Biological Chemistry*, *291*(41), 21596-21606.

84. Tremblay, J. M., Kuo, C. L., Abeijon, C., Sepulveda, J., Oyler, G., Hu, X., ... & Shoemaker, C. B. (2010). Camelid single domain antibodies (VHHs) as neuronal cell intrabody binding agents and inhibitors of Clostridium botulinum neurotoxin (BoNT) proteases. *Toxicon*, *56*(6), 990-998.

85. Campellone, K. G., Giese, N., Tipper, O. J., & Leong, J. M. (2002). A tyrosine-phosphorylated 12-amino-acid sequence of enteropathogenic Escherichia coli Tir binds the host adaptor protein Nck and is required for Nck localization to actin pedestals. *Molecular microbiology*, *43*(5), 1227-1241.

86. Kan, A., Birnbaum, D. P., Praveschotinunt, P., & Joshi, N. S. (2019). Congo red fluorescence for rapid in situ characterization of synthetic curli systems. *Applied and environmental microbiology*, *85*(13).

87. Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods*, *6*(5), 343-345.

88. Vertzoni, M., Diakidou, A., Chatzilias, M., Söderlind, E., Abrahamsson, B., Dressman, J. B., & Reppas, C. (2010). Biorelevant media to simulate fluids in the ascending colon of humans and their usefulness in predicting intracolonic drug solubility. *Pharmaceutical research*, *27*(10), 2187-2196.

89. Yang, Z., Schmidt, D., Liu, W., Li, S., Shi, L., Sheng, J., ... & Feng, H. (2014). A novel multivalent, single-domain antibody targeting TcdA and TcdB prevents fulminant Clostridium difficile infection in mice. *The Journal of infectious diseases*, *210*(6), 964-972.

Chapter 4

Conclusions

4.1. Summary

In this dissertation, I presented several results related to the use of engineered *E. coli* Nissle 1917 (EcN) for research and clinical applications. As an important chassis organism for synthetic biology endeavors, EcN remains a relevant microbe for an increasing number of researchers, and is currently being developed for clinical use[1-5]. Novel tools and applications for EcN engineering are therefore necessary for living therapeutics and diagnostics to reach their full potential.

In chapter 2, we described the development of a new plasmid system for heterologous gene expression in EcN – engineered pMUT plasmids. Based on the two plasmids found natively in wild-type EcN, pMUT1 and pMUT2, these engineered pMUTs were demonstrated as versatile vectors for a diverse set of genetic constructs. Importantly, the pMUT plasmids do not require the use of antibiotics for retention. We have demonstrated that this property extends to their engineered counterparts, both *in vitro* and *in vivo*, even when utilized to express an exogenous operon in the harsh environment of the gastrointestinal tract. In addition to the primary results described above, our efforts to remove the native copies of the pMUT plasmids led us to expand upon an existing plasmid curing system, and in the process highlighted the importance of the RelE-RelB toxin-antitoxin system for pMUT2 retention.

Chapter 3 proposed a novel therapeutic use of the BIND platform previously developed in the Joshi Laboratory[6], and detailed its testing *in vitro*. By fusing CsgA, the major curli subunit, to VHH domains, we were able to obtain a functional, EcN-tethered protein matrix able to specifically bind virulence factors of several enteric pathogens. These efforts, done in collaboration with the laboratories of Prof. John Leong and Prof. Charles Shoemaker, also

involved the generation of novel nanobody sequences, which were tested for their ability to bind targets associated with the cell surface of several attaching and effacing pathogens. While some existing work has utilized various systems to display VHH domains on *E. coli* [7,8], our results in chapters 2 and 3 constitute, to our knowledge, the first examples of CsgA-VHH fusions. Chapter 2 included the use of CsgA-NbGFP to demonstrate the feasibility of functional CsgA-VHH production by EcN, whereas the work described in chapter 3 greatly expanded the range of VHHs explored, and put forward an appropriate therapeutic application. Nanobodies and curli fibers have both been proposed as components for use in engineered living therapeutics. When used in combination, these elements complement one another: the curli fibers providing a self-assembling scaffold for multivalent display, while the VHHs provide binding functionality. We hope our work lays the foundation for further exploration and development of this unique system.

4.2. Limitations

While the engineered pMUT plasmids described in chapter 2 can readily be utilized as research tools, questions remain regarding their suitability for clinical use. In engineered microbial strains developed as therapeutics, genomic integration of heterologous genes is typically favored over plasmid-based approaches. This is in part due to the antibiotics required for plasmid maintenance, but also due to concerns regarding biocontainment and horizontal gene transfer. Therefore, while the use of antibiotics for plasmid retention does not appear to be of concern for the engineered pMUTs, the latter aspect requires further consideration.

Although neither of the pMUT plasmids contain the full suite of genetic components required for conjugation[9], our data cannot rule out horizontal gene transfer from these plasmids, nor can it be used to estimate its extent. It should be noted, however, that the genes found natively on the pMUT plasmids are unlikely to pose a major risk to patients, considering their presence in wild-type EcN, which has exhibited an excellent track record of safe use over multiple decades. The question then becomes whether this assumption of safety can be extended to the heterologous genes a pMUT-based therapeutic strain would be engineered to contain, a question that may need to be considered on a case-by-case basis.

An additional consideration involves the relationship between EcN density in the gut and the level of heterologous gene expression. As we have seen in our *in vivo* studies, the density of EcN in fecal samples continued to drop over the course of the experiment, likely due to the recovery of the native mouse microbiome in the absence of antibiotic administration. EcN is not native to mice and is likely to fare better in humans. Nevertheless, regardless of host species, continuous expression of heterologous proteins by EcN in the gut environment is likely to impose a high metabolic burden and impair the fitness of the engineered strain. Gene expression must therefore be tuned to strike a balance between population size and expression strength, in order to maximize therapeutic output. This trade-off applies to virtually all engineered living therapeutics, though it may be of particular relevance for strains bearing the engineered pMUTs due to the higher copy number associated with plasmid-based gene expression.

The curli-based pathogen sequestration approach described in chapter 3 presents a novel, versatile platform intended for the treatment or prophylaxis of infectious enteric disease. However, despite the promising *in vitro* results, assumption of *in vivo* therapeutic efficacy would be premature. In particular, this pathogen sequestration strategy relies on the

hypothesis that pathogens bound by the sequestrant matrix are less likely to interact with the host cells prior to being cleared from the gastrointestinal tract. While this hypothesis appears reasonable in light of the results presented in chapter 3, further work is required to test this for each target pathogen. Enteric pathogens vary in their infective dose and mechanism of pathogenicity[10], as do the binding characteristics of different curli-VHH matrices.

4.3. Future directions

While chapter 2 included extensive characterization of the performance of the engineered pMUT plasmids as expression vectors, additional features of the system remain to be explored. For instance, many genetic circuits utilized in synthetic biology projects require simultaneous use of multiple plasmids[11,12]. Future work could recapitulate such circuits by utilizing modified versions of both pMUT plasmids within the same EcN strain.

Additional research into the nature of the native pMUT plasmids could also benefit their further utilization and modification. Continuing to elucidate their relationship with EcN and the full scope of cellular mechanisms involved in their retention would inform the design of not only future pMUT variations, but other genetic systems as well. Such efforts would also shed light on the biology of EcN, itself a clinically important microorganism.

As mentioned above, a potential limitation of the pMUT plasmids in clinical use involves concerns regarding horizontal gene transfer. Further studies are therefore needed to evaluate the extent of this phenomenon. Notably, any outcome here would be of potential clinical interest. An absence of gene flow between pMUT-bearing bacteria and the native microbiome may support the system's applicability for use in humans. The presence of

substantial gene transfer, however, could pave the way for the use of horizontal gene transfer as a novel therapeutic strategy, as previously proposed by Sheth and coworkers[13].

As for the curli-VHH system developed and presented in chapter 3, multiple future research directions remain to be explored, chief among which is testing in an animal model. Notably, the selection and development of such model systems is a nontrivial process, with several factors to consider. First, both the pathogen and the engineered EcN strain must be administered, and must also maintain relatively reproducible and stable populations for an experimentally relevant duration of time. This would require optimization of bacterial doses and administration schedule. Additionally, the pathogen must be able to produce a disease phenotype in the presence of EcN, to allow for evaluation of curli-VHH efficacy irrespective of its chassis organism. Such considerations must take into account the pathogen biology, its interaction with the host, the identity of the VHH target associated with the pathogen, host-EcN relationship, and more.

Lastly, a more detailed investigation of the structure of the curli-VHH network would further improve our understanding of the system, and would subsequently inform its future design. As mentioned in chapter 3, both binding capacity and fiber morphology varied between different VHH domains. Future insights into the relationship between VHH sequence, curli-VHH matrix structure, and functionality of the appended binding domains would allow for the selection of VHHs that yield improved performance when fused to CsgA, thereby increasing the system's therapeutic potential.

4.4. References

1. Ou, B., Yang, Y., Tham, W. L., Chen, L., Guo, J., & Zhu, G. (2016). Genetic engineering of probiotic Escherichia coli Nissle 1917 for clinical application. *Applied microbiology and biotechnology*, *100*(20), 8693-8699.

2. Pedrolli, D. B., Ribeiro, N. V., Squizato, P. N., de Jesus, V. N., Cozetto, D. A., Tuma, R. B., ... & Cerri, M. O. (2019). Engineering microbial living therapeutics: the synthetic biology toolbox. *Trends in biotechnology*, *37*(1), 100-115.

3. Isabella, V. M., Ha, B. N., Castillo, M. J., Lubkowicz, D. J., Rowe, S. E., Millet, Y. A., ... & Falb, D. (2018). Development of a synthetic live bacterial therapeutic for the human metabolic disease phenylketonuria. *Nature biotechnology*, *36*(9), 857-864.

4. Kurtz, C. B., Millet, Y. A., Puurunen, M. K., Perreault, M., Charbonneau, M. R., Isabella, V. M., ... & Miller, P. F. (2019). An engineered E. coli Nissle improves hyperammonemia and survival in mice and shows dose-dependent exposure in healthy humans. *Science translational medicine*, *11*(475).

5. National Library of Medicine (U.S.). (2018, April 17 – 2019, June 21). *Safety and Tolerability of SYNB1618 in Healthy Adult Volunteers and Adult Subjects With Phenylketonuria (PKU).* Identifier NCT03516487. https://clinicaltrials.gov/ct2/show/NCT03516487

6. Nguyen, P. Q., Botyanszki, Z., Tay, P. K. R., & Joshi, N. S. (2014). Programmable biofilm-based materials from engineered curli nanofibres. *Nature communications*, *5*(1), 1-10.

7. Piñero-Lambea, C., Bodelón, G., Fernández-Periáñez, R., Cuesta, A. M., Álvarez-Vallina, L., & Fernández, L. A. (2015). Programming controlled adhesion of E. coli to target surfaces, cells, and tumors with synthetic adhesins. *ACS synthetic biology*, *4*(4), 463-473.

8. Klein, Á., Kovács, M., Muskotál, A., Jankovics, H., Tóth, B., Pósfai, M., & Vonderviszt, F. (2018). Nanobody-displaying flagellar nanotubes. *Scientific reports*, *8*(1), 1-9.

9. Sonnenborn, U., & Schulze, J. (2009). The non-pathogenic Escherichia coli strain Nissle 1917–features of a versatile probiotic. *Microbial Ecology in Health and Disease*, *21*(3-4), 122-158.

10. Greig, J. D., Todd, E. C., Bartleson, C., & Michaels, B. (2010, March). Infective doses and pathogen carriage. In *2010 Food Safety Education Conference*.

11. Elowitz, M. B., & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, *403*(6767), 335-338.

12. Nandagopal, N., & Elowitz, M. B. (2011). Synthetic biology: integrated gene circuits. *science*, *333*(6047), 1244-1248.

13. Sheth, R. U., Cabral, V., Chen, S. P., & Wang, H. H. (2016). Manipulating bacterial communities by in situ microbiome engineering. *Trends in Genetics*, *32*(4), 189-200.

Appendix A

Mathematical model of probiotic and pathogen population dynamics *in vivo*

A.1. Rationale and motivation

The curli-based sequestration approach presented in chapter 3 entails the simultaneous presence of two microbial species – an engineered probiotic and a pathogen – in the gastrointestinal tract of a host. To interfere with pathogen-host interactions and hasten pathogen clearance, the engineered probiotic – in this case, *E. coli* Nissle 1917 (EcN) – produces a protein matrix functionalized with pathogen-binding domains. However, as discussed in chapter 4, the *in vitro* results of chapter 3 would require validation *in vivo* to demonstrate their efficacy, which would necessitate the development of an appropriate experimental model system. At the time of writing, such animal models are not available to us. At the suggestion of my graduate committee, I developed a mathematical model for the populations of both microbial species, in an attempt to estimate the possible range of behaviors the curli-based sequestration system would exhibit *in vivo*, based on data available in the literature. Considering both the novelty and the complexity of the pathogen sequestration approach, such estimates alone are not sufficient to predict the therapeutic outcome. However, if applied to future *in vivo* data, this model may assist in experimental design. Further, this mathematical model may also be of use to researchers designing and characterizing other systems in which a pathogen and a probiotic are co-administered, potentially allowing for the quantification of the probiotic's inhibitory effect on the pathogen and facilitating comparison between different strains.

A.2. Mathematical framework

Previous efforts to model microbiome population dynamics have utilized a variety of mathematical formulations[1-3]. As a starting point for the model described here, I used a formulation based on the generalized Lotka-Volterra equations[4], similar to one previously described by Gibson and Gerber[5]. However, the model was simplified considerably due to its intended application: while such models are designed for inferring population dynamics of multiple species on a microbiome-wide scale, the work described here tracks only two bacterial populations – pathogen and probiotic.

In light of these considerations, the following set of ordinary differential equations was used:

$$\frac{dN}{dt} = f + a_n N + b_{np} NP \quad (1)$$

$$\frac{dP}{dt} = a_p P + a_{pp} P^2 + b_{pn} PN \quad (2)$$

The parameters and their interpretations are summarized in Table A.1 below.

Table A.1: Model parameters

Parameter	Meaning	Units
N	*E. coli* Nissle 1917 (EcN) population	[CFU/g]
P	Pathogen population	[CFU/g]
f	Effective EcN feeding rate	[CFU/(g*day)]
a_n	EcN growth rate	[1/day]
b_{np}	EcN-pathogen interaction term (EcN equation)	[g/(CFU*day)]
a_p	Pathogen growth rate	[1/day]
a_{pp}	Pathogen self-interaction term	[g/(CFU*day)]
b_{pn}	EcN-pathogen interaction term (pathogen equation)	[g/(CFU*day)]

The differences in form between equations (1) and (2) are related to the differing strain characteristics, as well as the experimental setup. To start, the experimental design recapitulated here involves regular feeding of EcN, and is expressed by the term f in equation (2). In contrast, such a term does not appear in equation (2), as pathogen exposure only occurs once, and can therefore be captured by the initial conditions. Furthermore, the quadratic "self-interaction" term $a_{pp}P^2$ appears only in the pathogen equation. An analogous term was not included in the EcN equation, as it was not required to capture EcN behavior – a constant, steady-state bacterial density during feeding, and a roughly exponential decay once feeding is halted. The pathogen self-interaction term, however, was required to achieve a nonzero steady-state for pathogen density in leu of a feeding term.

A.3. Collection of literature data and model fitting

Next, to obtain parameter estimates for further modeling, literature data was used. To match the experimental system the mathematical model aims to recapitulate, any study used must feature feeding of EcN in addition to a pathogenic *E. coli* or *Citrobacter rodentium* (CR) strain. Furthermore, both probiotic and pathogen population density must be estimated from fecal samples over several days. Unfortunately, such data is scarce, and only a single such study was utilized[6]. The dataset was therefore supplemented with three additional studies in which only one of the two strains was used – either EcN or a pathogen (in this case, CR), one of which is the one described in chapter 2. These data were used to get a baseline estimate for the parameters associated with each strain. Unfortunately, in addition to the paucity of available data, the studies also varied in pathogen strain and host organism, and in addition provided sparse, high-variance data. However, these data were

nevertheless suitable for the purpose of fitting the model and arriving at rough parameter estimates. The data used is summarized below in Table A.2.

Table A.2: Microbial population data used for model fitting

Source	Microbe	Host	Feeding start day	Feeding stop day	Dose [CFU]	Day	Density [CFU/g]
This dissertation (chapter 2)	EcN	Mouse	0	4	$5*10^7$	1	$6.96*10^7$
						2	$1.77*10^9$
						3	$1.96*10^9$
						4	$7.50*10^8$
					0	5	$1.90*10^8$
						6	$2.02*10^7$
						7	$4.05*10^6$
Kurtz *et al.*[7]	EcN	Cynomolgus monkey	0	28	$1*10^9$	2	$3.21*10^5$
						21	$1.06*10^5$
					0	30	$1.82*10^4$
						35	$1.25*10^3$
					$1*10^{12}$	2	$4.59*10^5$
						21	$1.45*10^7$
					0	30	$1.42*10^5$
						35	$9.39*10^3$
		Human	0	28	$2.37*10^{11}$	4	$6.00*10^4$
						6	$1.50*10^5$
						14	$7.72*10^4$
						20	$1.21*10^5$
						28	$6.07*10^4$
Flowers *et al.*[8]	CR(ΦStx2dact)	Mouse	1	1	$3*10^4$	1	$1.06*10^3$
						5	$7.92*10^6$
						7	$2.20*10^8$
						9	$6.85*10^8$
			1	1	$3.6*10^5$	1	$9.79*10^2$
						5	$1.02*10^8$
						7	$1.72*10^9$
						9	$3.13*10^9$
			1	1	$7.5*10^7$	1	$5.84*10^4$
						5	$3.28*10^8$
						7	$3.51*10^8$
			1	1	$3.5*10^8$	1	$2.62*10^5$
						5	$2.58*10^8$
						7	$3.96*10^8$
Sonnenborn *et al.*[6]	*E. coli* 542/88	Gnotobiotic piglet	0	0	$1*10^8$	0	$8.96*10^7$
						2	$9.65*10^9$
						4	$1.81*10^0$
						6	$1.44*10^0$
						8	$7.67*10^8$
						10	$6.24*10^7$
						12	$1.29*10^0$
	EcN		6	18	$2*10^8$	6	$9.76*10^7$
						8	$3.16*10^0$
						10	$1.84*10^0$
						12	$2.76*10^0$
						18	$3.45*10^0$

Next, a Python program was designed to fit the data and obtain estimates for model parameters for each dataset used. A grid search strategy was utilized to explore the parameter space. For each set of parameter values, the system of differential equations was solved numerically using the Scipy odeint solver. For EcN, the initial value was set to 0, while the initial pathogen density was taken as the first available datapoint. A score was calculated by taking the sum of squares error of the log-transformed model predictions at the available datapoints, and the combination of parameters yielding the minimal error was chosen, yielding parameter estimates.

This workflow was first used on single strain data, starting with EcN. In the absence of pathogen, only 2 of the 6 parameters can take on nonzero values: the feeding term f, and the EcN rate constant a_n. Similarly, the model was then applied to the pathogen-only datasets, with the only 2 nonzero parameters being a_p and a_{pp}. The plots of the EcN-only and pathogen-only models with the optimized parameters can be found in Figure A.1 and A.2, respectively, and the parameter values are summarized in Table A.3.

Table A.3: Summary of optimal parameter values

Dataset	Host	f	a_n	b_{np}	a_p	a_{pp}	b_{pn}
Chapter 2	Mouse	$1.259*10^{9}$	$-1.7*10^{0}$	N/A	N/A	N/A	N/A
Kurtz	Cynomolgus monkey	$1.259*10^{5}$	$-7.5*10^{-1}$	N/A	N/A	N/A	N/A
		$1.585*10^{6}$	$-8.0*10^{-1}$	N/A	N/A	N/A	N/A
	Human	$5.012*10^{4}$	$-5.5*10^{-1}$	N/A	N/A	N/A	N/A
Flowers	Mouse	N/A	N/A	N/A	$2.1*10^{0}$	$-3.162*10^{-9}$	N/A
		N/A	N/A	N/A	$2.9*10^{0}$	$-1.122*10^{-9}$	N/A
		N/A	N/A	N/A	$3.1*10^{0}$	$-8.913*10^{-9}$	N/A
		N/A	N/A	N/A	$1.9*10^{0}$	$-4.467*10^{-9}$	N/A
Sonnenborn	Gnotobiotic piglet	N/A	N/A	N/A	$2.8*10^{0}$	$-1.783*10^{-10}$	N/A
		$6.310*10^{9}$	$-7.924*10^{-2}$	0	$2.8*10^{0}$	$-1.783*10^{-10}$	$-3.594*10^{-10}$

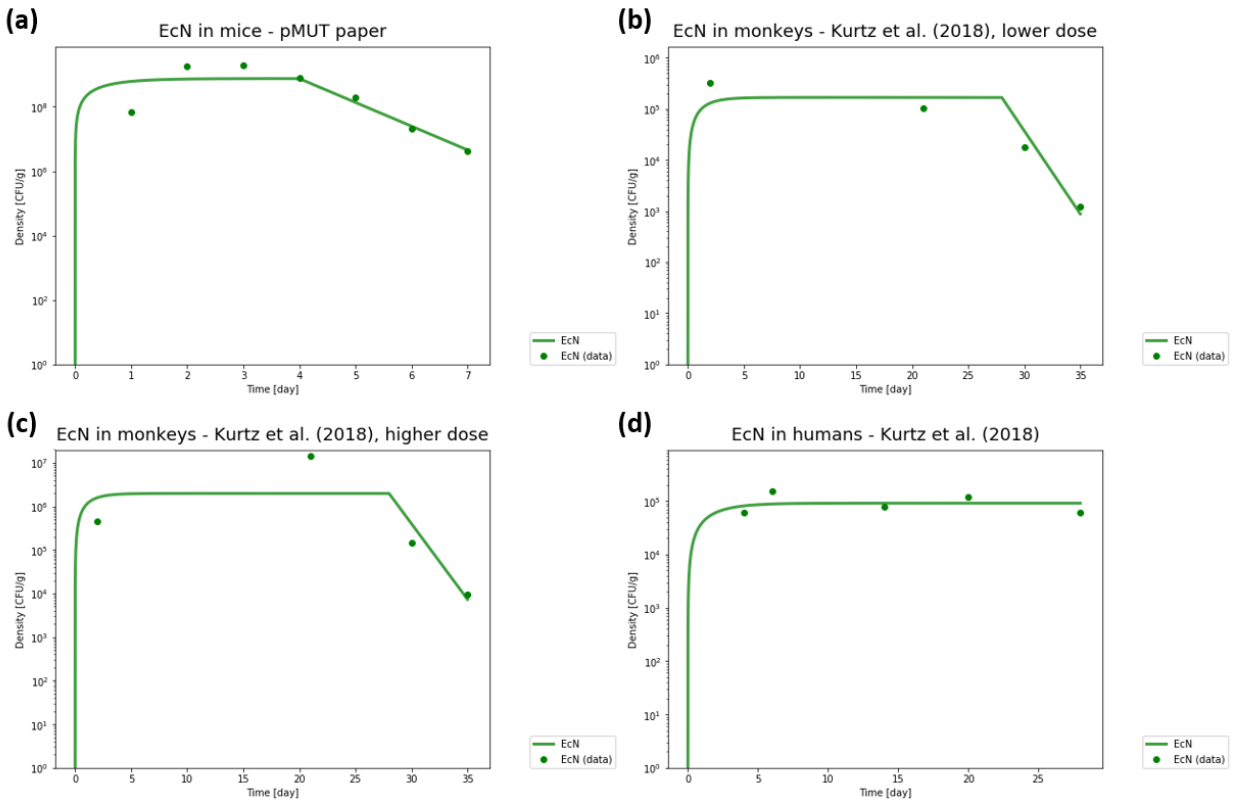


Figure A.1: Model fitting to available EcN-only data. (a) EcN in mice, chapter 2 (b) EcN in cynomolgus monkeys, lower dose, Kurtz *et al.* (c) EcN in cynomolgus monkeys, higher dose, Kurtz *et al.* (d) EcN in humans, Kurtz *et al.* Parameter values correspond to Table A.3.

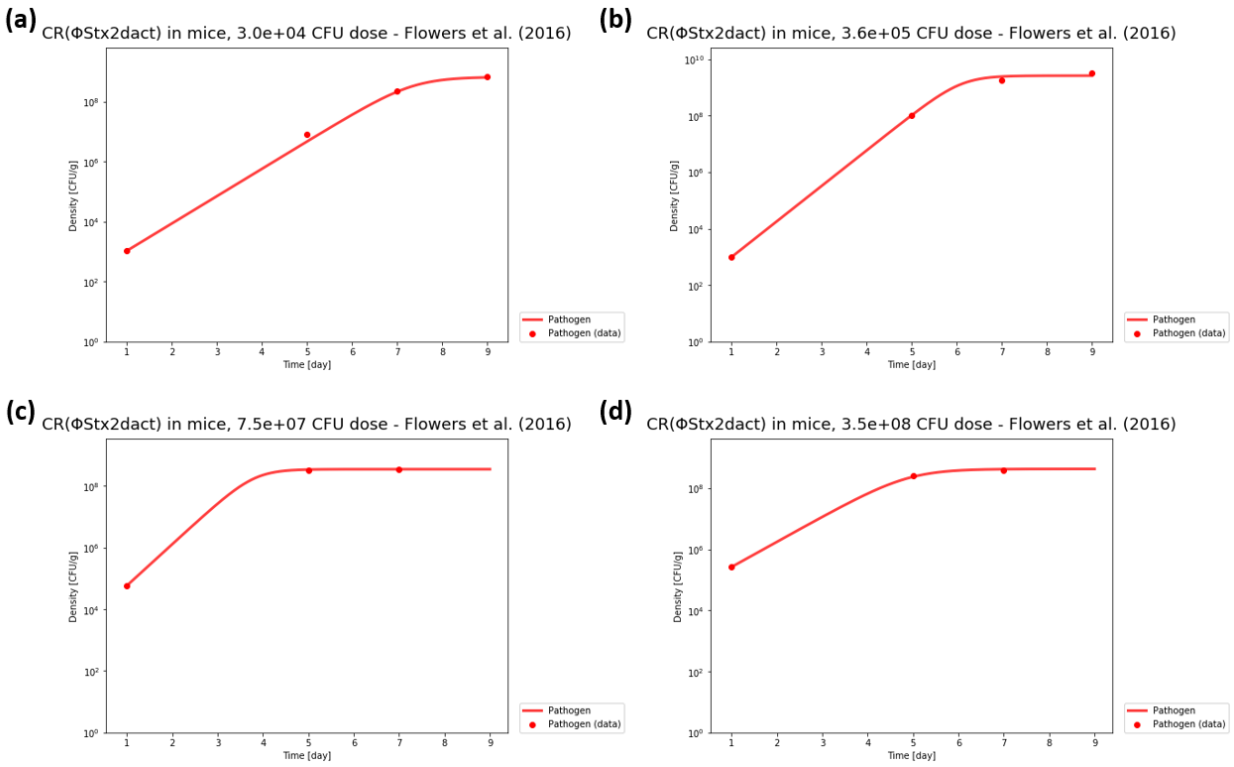


Figure A.2: Model fitting to available pathogen-only data. *C. rodentium* model in mice developed by Flowers *et al.* Doses of $3*10^4$ (a), $3.6*10^5$ (b), $7.5*10^7$ (c), and $3.5*10^8$ colony forming units (d). Parameter values correspond to Table A.3.

The EcN parameters obtained from the probiotic-only data generally matched expected trends. Although the feeding rate varied between datasets, this was not surprising considering the different doses and administration schedules applied in each case. In contrast, the rate constant a_n remained within a relatively limited range for all conditions, between -1.7 day^{-1} to -0.55 day^{-1}. Furthermore, the higher-magnitude value of -1.7 day^{-1}, which corresponds to more rapid EcN clearance, was obtained in mice, whereas the slowest clearance was exhibited in the human data. This is the trend we would expect given that EcN is a human isolate. The pathogen-only data was relatively consistent as well, though this was perhaps to be expected, as all four datasets used were taken from the same study

with the same model system. Pathogen rate constants ranged from 1.9 to 3.1 day^{-1} for a_p, and -1.2*10^{-9} to -8.1*10^{-9} g CFU^{-1} day^{-1} for a_{pp}.

To estimate the interaction terms b_{np} and b_{pn}, the study described by Sonnenborn and Schulze was used. Since EcN administration began 6 days after pathogen inoculation, the first timepoints could be utilized to obtain pathogen-only parameters (Figure A.3a). These parameters were then used when fitting the full model (Figure A.3b). The EcN parameters obtained here differed substantially from those previously obtained, perhaps due to the fact this work was performed in a gnotobiotic model. Nevertheless, this provides us with a starting point for further modeling. Interestingly, the value obtained for the parameter b_{np} was 0, suggesting pathogen effect on EcN population was negligible in this study.

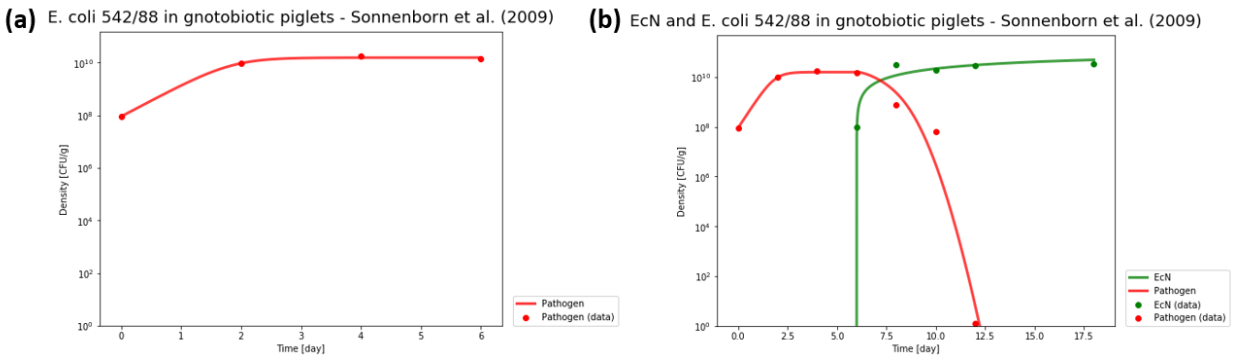


Figure A.3: Model fitting to EcN-pathogen data. (a) Data from days 0 to 6 were first used to estimate pathogen-only parameters. (b) Full model, including EcN and interaction terms.

A.4. Simulated experiments

The parameters obtained from available data were used to run several simulations of EcN-pathogen co-administration experiments. For pathogen parameters, the data for *C. rodentium* was used, as it was the most reliable. Values for a_p and a_{pp} were chosen from the range previously mentioned, with the values used being 2.5 day^{-1} and $-5*10^{-9}$ g CFU^{-1} day^{-1}, respectively. The initial value was set at $1*10^6$ colony forming units. Since the corresponding pathogen data was gathered in mice, the EcN parameter values used were the ones obtained in the mouse study from chapter 2 (values shown in the first row of Table A.3). In these simulations, the feeding of EcN begins 2 days after initial infection.

The simulations entailed three scenarios: (a) EcN is unaffected by pathogen population ($b_{np} = 0$); (b) EcN is moderately affected by pathogen population ($b_{np} = -1 \cdot 10^{-8}$); and (c) EcN is strongly affected by pathogen population ($b_{np} = -1 \cdot 10^{-7}$). Within each scenario, multiple values of b_{np} were examined, reflecting varying levels of EcN's ability to reduce pathogen populations. The results are plotted in Figure A.4.

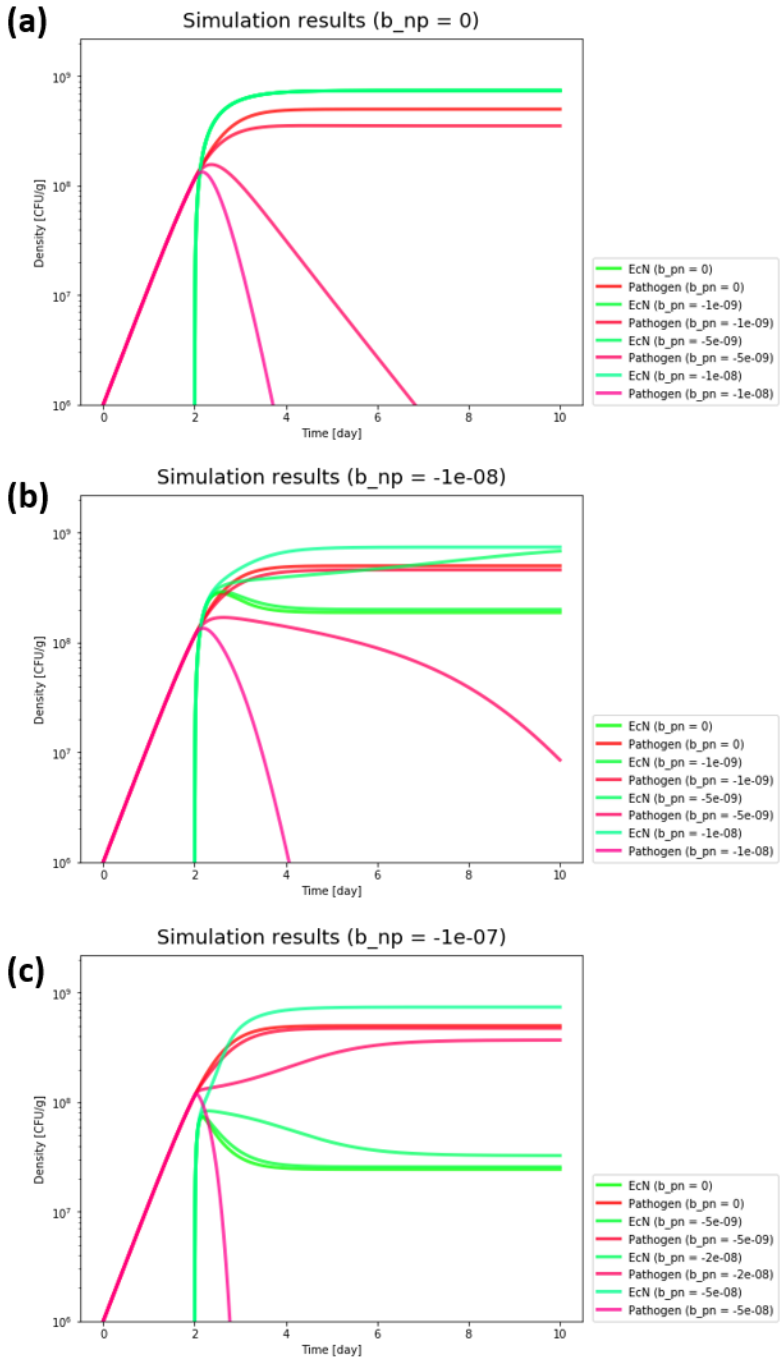


Figure A.4: Results of simulated experiments. The value of the EcN-pathogen interaction parameter b_{np} was set to (a) 0, (b) $-1*10^{-8}$, or (c) $-1*10^{-8}$ g CFU^{-1} day^{-1}.

Overall, in scenarios (a) and (b), where EcN is only weakly influenced by pathogen population, a b_{pn} value of approximately -5*10^{-9} g CFU^{-1} day^{-1} or higher in magnitude was necessary to see a reduction in pathogen population. This value is an order of magnitude higher than the one observed in the gnotobiotic piglet study, potentially suggesting that wild-type EcN is unlikely to ameliorate *C. rodentium* infection in mice. Furthermore, as the requirements on EcN performance became even more demanding in scenario (c), the threshold b_{pn} further shifted to a minimum of approximately -5*10^{-8} g CFU^{-1} day^{-1}, an additional order of magnitude higher. Due to the stronger interaction with the pathogen, EcN population also appeared to reach a lower steady state, except for in cases where it was able to substantially reduce pathogen population, exemplifying inter-species competition.

A.5. Conclusions

The model presented here describes the population-level behavior of EcN and a pathogen species simultaneously inhabiting a host organism. Although it was successful in fitting some existing experimental data and recapitulating overall population dynamics, the paucity of appropriate data currently presents a hurdle for further use of this model. Through simulation of experiments using estimated parameters, the model was also able to provide some insight into the relationship between probiotic and pathogen populations, and how population size was as important as interaction strength in dictating dynamics.

Drawing on these insights to inform design of engineered EcN systems, this latter point highlights an important consideration: while production of therapeutic proteins by the engineered probiotic can improve its ability to combat the infectious agent (i.e., increasing the magnitude of b_{pn}), this protein expression can also reduce its own population by

imposing a high metabolic burden, thereby potentially canceling out the benefit gained. Therefore, strain optimization is required for successful therapeutic output. This trade-off is well appreciated within the field of engineered living therapeutics, and it should therefore come as no surprise that it is present in the results presented here.

While this model was ultimately only partially successful in providing insight into the feasibility of curli-based pathogen sequestration *in vivo* (or of any other engineered probiotic effort), it may nevertheless have additional merit. By fitting this model to other pathogen-probiotic population data and estimating model parameter, it may be used to compare between probiotic formulations in a quantitative manner. Further, as future work generates relevant *in vivo* data for the testing of curli-based pathogen sequestration, we hope this model can be used to facilitate experimental planning, quantitatively informing dosing and bacterial administration schedule.

A.6. References

1. Bucci, V., & Xavier, J. B. (2014). Towards predictive models of the human gut microbiome. *Journal of molecular biology*, *426*(23), 3907-3916.
2. Roughgarden, J. (2019). Holobiont evolution: mathematical model with vertical vs. horizontal microbiome transmission. *Philosophy, Theory, and Practice in Biology*, *12*(002).
3. Dong, Y., Takeuchi, Y., & Nakaoka, S. (2018). A mathematical model of multiple delayed feedback control system of the gut microbiota—Antibiotics injection controlled by measured metagenomic data. *Nonlinear Analysis: Real World Applications*, *43*, 1-17.
4. Stein, R. R., Bucci, V., Toussaint, N. C., Buffie, C. G., Rätsch, G., Pamer, E. G., ... & Xavier, J. B. (2013). Ecological modeling from time-series inference: insight into dynamics and stability of intestinal microbiota. *PLoS Comput Biol*, *9*(12), e1003388.
5. Gibson, T., & Gerber, G. (2018, July). Robust and scalable models of microbiome dynamics. In *International Conference on Machine Learning* (pp. 1763-1772). PMLR.
6. Sonnenborn, U., & Schulze, J. (2009). The non-pathogenic Escherichia coli strain Nissle 1917–features of a versatile probiotic. *Microbial Ecology in Health and Disease*, *21*(3-4), 122-158.
7. Kurtz, C., Denney, W. S., Blankstein, L., Guilmain, S. E., Machinani, S., Kotula, J., ... & Brennan, A. M. (2018). Translational Development of Microbiome-Based Therapeutics: Kinetics of E. coli Nissle and Engineered Strains in Humans and Nonhuman Primates. *Clinical and translational science*, *11*(2), 200-207.
8. Flowers, L. J., Bou Ghanem, E. N., & Leong, J. M. (2016). Synchronous disease kinetics in a murine model for enterohemorrhagic E. coli infection using food-borne inoculation. *Frontiers in cellular and infection microbiology*, *6*, 138.

Appendix B

Chapter 2 supplementary information

Supplementary Table B1. Primers used to assess pMUTs in colony PCR, as well as the primers used to insert the recombinant cassettes onto the pMUT plasmids to make the engineered versions. In all cases, the capitalized sequences indicate homology to the pMUT target. Primer pairs in italic could not produce an amplicon with EcN DNA as a template.

Primer Name	Sequence	Target
Muta 5	AACTGTGAAGCGATGAACCC	pMUT1
Muta 6	GGACTGTTCAGAGAGCTATC	pMUT1
Muta 7	GACCAAGCGATAACCGGATG	pMUT2
Muta 8	GTGAGATGATGGCCACGATT	pMUT2
pM1S2chk_F	GAATAGGGTGACACTGGCGCC	pMUT1
pM1S2chk_R	CCAGATGGCATTGTAACAGACTTCTC	pMUT1
pM1S3chk_F	CGCATCCTTCCTGTTTTTCCGG	pMUT1
pM1S3chk_R	GTCTTGGTAGCCCTGCTTCTGG	pMUT1
pM2S2chk_F	AGTTTCGCACCCAAAGTGCG	pMUT2
pM2S2chk_R	GACAAAACAACCTATATCAGATAACAGC	pMUT2
pM2S3chk_F	GATAAAACTATCAACTCACCGTCTTG	pMUT2
pM2S3chk_R	GCCGTTGGTCTTTACTGATTTAAG	pMUT2
M1_1F	aaaccttgcttcttcgcggtGAATAGGGTGACACTGGCG	pMUT1
M1_1R	accgcattctagatttagggATATGAATATACCATATAATATATACTTTAAATATTTTGG	pMUT1
M1_1F2	aaaccttgcttcttcgcggtGAATAGGGTGACACTGGCGCCATTATTGTG	pMUT1
M1_1R2	accgcattctagatttagggATATGAATATACCATATAATATATACTTTAAATATTTTGGGGCTTAG	pMUT1
M1_2F	aaaccttgcttcttcgcggtGGAGTTAGCGATATGAAAACCGAACAACG	pMUT1
M1_2R	accgcattctagatttagggGTATCTAATTCAGGCAGGAAAAAATCTTTTCC	pMUT1
M1_3F	aaaccttgcttcttcgcggtGTTTCAGTGGTGCGTACAATTAAG	pMUT1
M1_3R	accgcattctagatttagggGCGCTGAACGCGATTCTG	pMUT1
M2_1F	aaaccttgcttcttcgcggtCCACTAAGTTACACCTCAACAACG	pMUT2
M2_1R	accgcattctagatttagggCAGAAAAAACAAAGCCCCG	pMUT2
M2_1F2	aaaccttgcttcttcgcggtCCACTAAGTTACACCTCAACAACGG	pMUT2
M2_1R2	accgcattctagatttagggCAGAAAAAACAAAGCCCCGAAATCATGC	pMUT2
M2_2F	aaaccttgcttcttcgcggtCCACTAAGTTACACCTCAACAACGG	pMUT2
M2_2R	accgcattctagatttagggCAGAAAAAACAAAGCCCCGAAATCATGC	pMUT2
M2_3F	aaaccttgcttcttcgcggtATTAAATAATGACAATGTTGGGTTG	pMUT2
M2_3R	accgcattctagatttagggGTTTCTGCCTATAAGATTACTTACAGTG	pMUT2

Supplementary Table B2. UNS sequences used in the engineered pMUT plasmid designs.

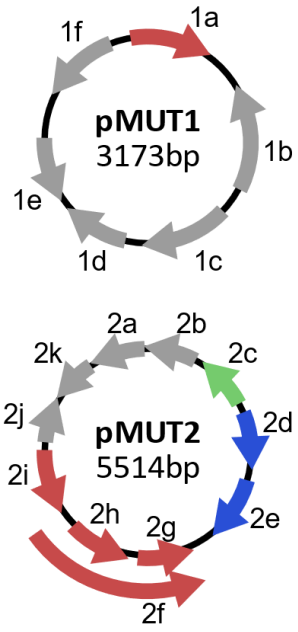
Name	Sequence
UNS0	GTTCCTTATCATCTGGCGAATCGGACCCACAAGAGCACTG
UNS1	CATTACTCGCATCCATTCTCAGGCTGTCTCGTCTCGTCTC
UNS2	GCTGGGAGTTCGTAGACGGAAACAAACGCAGAATCCAAGC
UNS3	GCACTGAAGGTCCTCAATCGCACTGGAAACATCAAGGTCG
UNS4	CTGACCTCCTGCCAGCAATAGTAAGACAACACGCAAAGTC
UNS5	GAGCCAACTCCCTTTACAACCTCACTCAAGTCCGTTAGAG
UNS6	CTCGTTCGCTGCCACCTAAGAATACTCTACGGTCACATAC
UNSX	CCAGGATACATAGATTACCACAACTCCGAGCCCTTCCACC

Supplementary Table B3. gRNA sequences used in the pFREE and pCryptDel plasmids. gRNAs 1-4 (starred) were found in the original pFREE sequence from Lauritsen *et al.* [1]. In each case, the gRNA name, sequence and target are shown.

gRNA Name	Sequence	Target
gRNA1*	atgaactagcgattagtcgctatgacttaa	Targets pSC101 origin
gRNA2*	aaccacactagagaacatactggctaaata	Targets pSC101 origin
gRNA3*	ggttggactcaagacgatagttaccggata	Targets ColE1-like origins except colA
gRNA4*	ggcgaaacccgacaggactataaagatacc	Targets ColE1-like origins including colA
gRNA5	ccgatttgatggctatcgttcgggatcgtc	Targets pMUT2 at RelE
gRNA6	aactgcaccctcttcgataaaacccgcaag	Targets pMUT2 at hypothetical protein
gRNA7	gctctcttttcaggagagtgatttacccga	Targets pMUT2 at relaxase
gRNA8	ttgattttgtagcagtcatgcagctctcgc	Targets pMUT2 at RelB
gRNA9	cttgaatttgatccccgagccctgaaggaa	Targets pMUT2 at RelB
gRNA10	gcccacactcaccatcaaaaaccccgagaa	Targets pMUT2 at relaxase

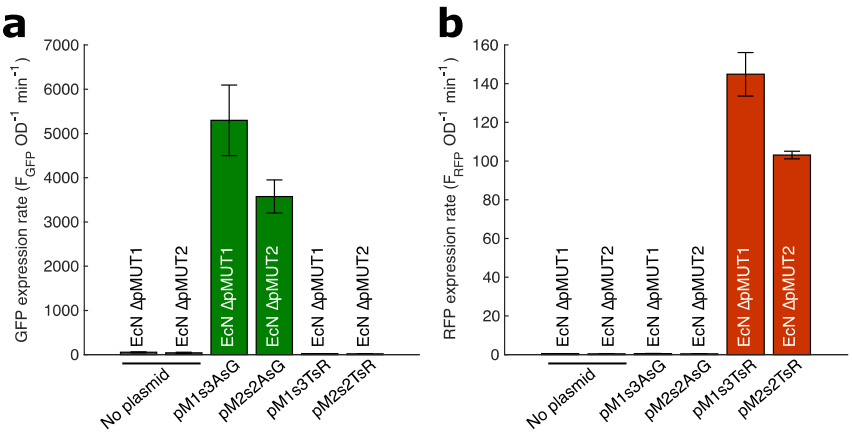
Supplementary Table B4. Variant TlpA* promoter sequences, and their promoter strengths, CsgA-FE36 production capacity, and CsgA-NbGFP production capacity at when bacteria were grown on 37°C and at 30°C. The leftmost column indicates the host engineered plasmid. Promoter strength was measured using constructs pM#s#AsR_TS*. Curli material was produced with plasmids containing the TlpA* variant driving cassette 'csg-Etag' (columns CsgA-FE36), or 'csg-Etag-NbGFP' (columns CsgA-NbGFP), and curli production was measured using the Congo Red method (CR) or by immunostaining with anti-Etag on a filter plate (ELISA).

Plasmid	Promoter variant	Promoter sequence	Promoter strength (GFP)		CsgA-FE36 production (CR)		CsgA-NbGFP production (CR)		CsgA-FE36 production (ELISA)		CsgA-NbGFP production (ELISA)	
			37°C	30°C	37°C	30°C	37°C	30°C	37°C	30°C	37°C	30°C
No plasmid control			65.5	58.0	1.00	1.00	1.00	1.00	0.10	0.10	0.08	0.09
pMUT1	A8	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataGGGaagc	588.2	123.2	1.02	1.07	0.95	1.01	0.24	0.09	0.11	0.07
pMUT1	A9	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataTGCaagc	721.4	137.9	1.06	1.11	1.11	1.08	0.27	0.08	0.29	0.09
pMUT1	C2	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataGGCaagc	372.7	121.3	1.00	1.07	0.95	1.00	0.11	0.08	0.07	0.08
pMUT1	C7	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataCAAaagc	2402.1	179.3	1.44	1.12	0.95	1.12	0.49	0.10	0.11	0.10
pMUT1	C8	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataGCTaagc	1478.6	158.1	1.03	1.08	1.27	1.08	0.09	0.08	0.51	0.10
pMUT1	D6	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataCGAaagc	1355.9	152.6	1.32	1.17			0.47	0.08		
pMUT1	D8	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataGAAaagc	618.1	128.0	1.03	1.14	1.43	1.11	0.13	0.09	0.84	0.11
pMUT2	E8	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataGAGaagc	963.6	172.2	1.41	1.10	1.32	1.08	0.35	0.08	0.54	0.10
pMUT2	E10	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataAAAaagc	1016.5	196.7	1.30	1.08	1.31	1.09	0.68	0.08	0.52	0.12
pMUT2	F7	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataGCAaagc	718.1	64.4	1.25	1.03	1.00	1.00	0.19	0.09	0.29	0.09

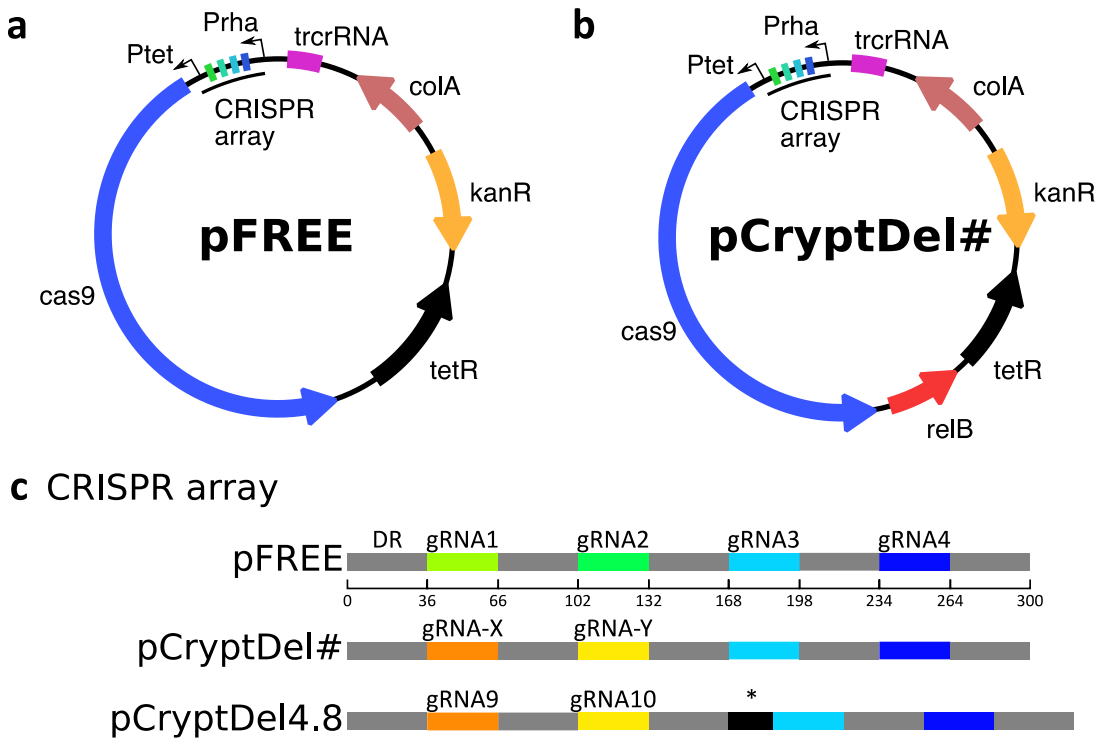


Predicted ORF	Gene name	Description	Notes
1a	*mobA*	Part of the relaxosome complex responsible for plasmid transfer during conjugation.	Part of plasmid conjugation machinery
1b	N/A	helix-turn-helix domain-containing protein	Putitative DNA-binding protein (MarR-like)
1c	N/A	Hypothetical protein	Multispecies, unknown function
1d	N/A	Hypothetical protein	Part of ColE1 plasmid replication origin, untranslated RNA
1e	N/A	Hypothetical protein	Part of ColE1 plasmid replication origin, untranslated RNA
1f	N/A	Rop family plasmid primer RNA-binding protein	Likely involved in plasmid copy number control
2a	N/A	Hypothetical protein	
2b	N/A	Hypothetical protein	
2c	*rep*	replicase	
2d	*RelB/Din*	Type II toxin-antitoxin system antitoxin	
2e	*RelE/StbE*	Addiction module toxin	
2f	*MobA*	Part of the relaxosome complex responsible for plasmid transfer during conjugation.	Part of plasmid conjugation machinery
2g	*MobD*	Plasmid mobilization protein	Part of plasmid conjugation machinery
2h	*MobB*	Essential to promote the specific transfer of the plasmid in the presence of conjugative plasmids.	Part of plasmid conjugation machinery
2i	*MobC*	Relaxosome protein	Part of plasmid conjugation machinery
2j	N/A	Hypothetical protein	
2k	N/A	DUF4868 domain-containing protein	Putitative bacteriocin?

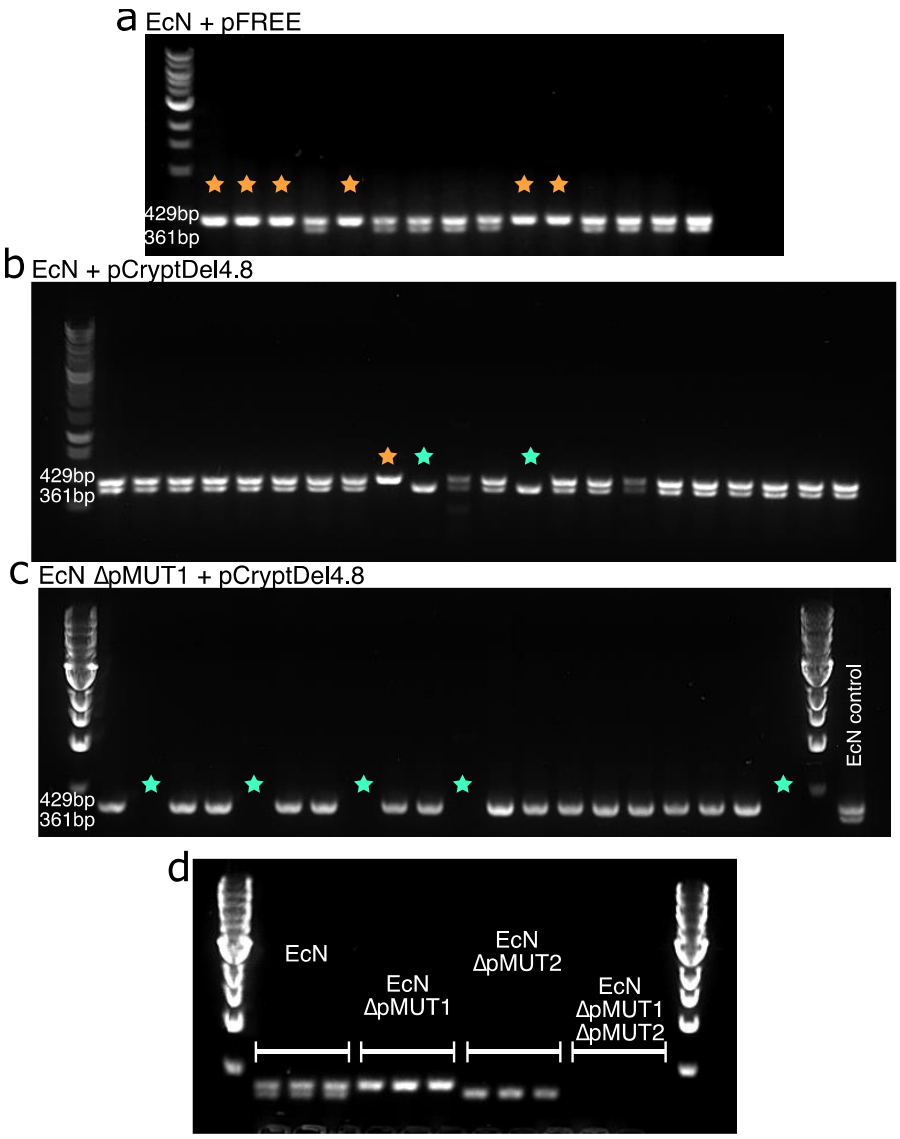
Supplementary Figure B1. Plasmid maps of pMUT1 and pMUT2 (left), alongside a table of annotations (right) for predicted ORFs in the native pMUT plasmids, with the 'Predicted ORF' referring to the labels on the plasmid maps.



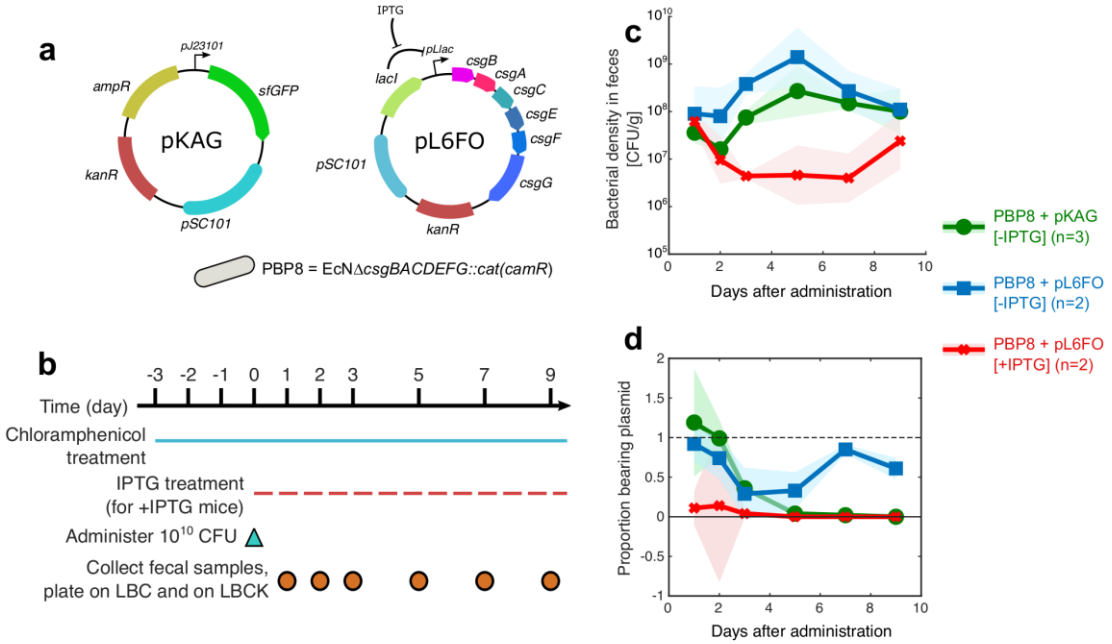
Supplementary Figure B2. Gene expression characterization from cassettes (a) 'AsG' and (b) 'TsR' on plasmids pM1s3 and pM2s2. In all cases, we used EcN without the relevant native pMUT plasmid. In both cases, pM1s3 plasmid backbone provides higher gene expression. The relative difference in recombinant protein expression strength between the engineered pMUT1 and pMUT2 plasmids is independent of the fluorescent protein used for characterization. Error bars show standard deviation from 8 replicates. Assays performed by Anton Kan and Siva Emani.



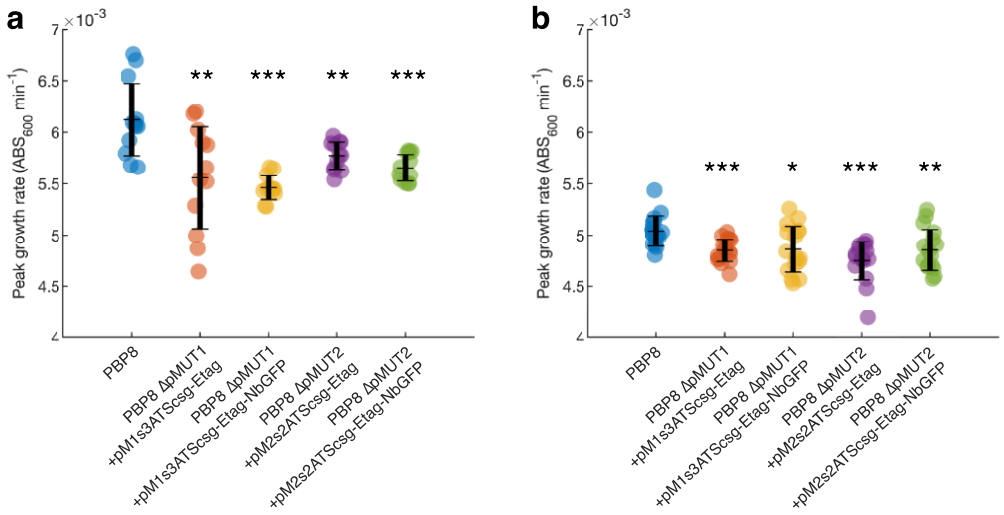
Supplementary Figure B3. Plasmid vectors to cure EcN cryptic plasmids. (a) Plasmid map of pFREE, showing ATC inducible Cas9 and rhamnose inducible CRISPR gRNA array, (b) plasmid map of the pCryptDel plasmid variants, which are based on pFREE with a modified CRISPR array and a *relB* anti-toxin gene. (c) A detailed look at the gRNA arrays. gRNA-X and gRNA-Y refer to variants (shown in Supplementary Table B3), where X and Y pair were either 5 and 6, 7 and 8, or 9 and 10. The final construct, pCryptDel4.8, contained gRNA9 and 10, but had a 34bp insertion in the region just upstream of gRNA3.



Supplementary Figure B4. Representative TAE agarose gels of colony PCR results following a typical pMUT curing process, removing native plasmids from EcN with (a) pFREE, (b) pCryptDel4.8, and (c) from EcN ΔpMUT1 with pCryptDel4.8. Panel (d) shows a gel for 3 replicate colonies each of unmodified EcN, and the pMUT knockout variants EcN ΔpMUT1, EcN ΔpMUT2, and EcN ΔpMUT1 ΔpMUT2. In all cases primers muta5, muta6, muta7 and muta8 were used, which result in a 429bp band in the presence of pMUT2, and a 361bp in the presence of pMUT1. In each case, an orange star shows a colony cured of pMUT1, and a blue star shows a colony cured of pMUT2. Assays performed by Anton Kan, Siva Emani, and Ilia Gelfat.



Supplementary Figure B5. (a) Synthetic plasmids pKAG, which constitutively expresses sfGFP, and pL6FO, which expresses the synthetic curli operon *csgBACEFG* with IPTG induction, were transformed into *E. coli* Nissle strain PBP8. (b) Administration and sampling schedule, with mice in all groups were treated with chloramphenicol to select for PBP8 cells from day -3 to the end of the experiment. On day 0, mice were administered with 10^{10} CFU of PBP8 transformed with either pKAG (n=3), or pL6FO (n=4), and half of the PBP8+pL6FO mice were given the IPTG inducer in their water. Fecal samples were collected regularly to detect PBP8 (chloramphenicol resistant) or PBP8 with plasmid (chloramphenicol and kanamycin resistant) by plating assays. (c) After administration, PBP8 cells were maintained in the mice for all conditions throughout the experiment, but (d) all plasmids suffered significant plasmid loss, particularly after day 2. Shaded areas show relative standard error for panel c and standard deviation for panel (d). Detailed assay methodology can be found in supplementary methods below. Assays performed by Pichet Praveschotinunt and Anton Kan.



Supplementary Figure B6. *In vitro* growth rates of bacterial strains used in the *in vivo* experiments grown at (a) 37°C and (b) 30°C. Growth rates were measured by measuring the absorbance at 600 nm of the bacterial cultures in a plate reader every 10 minutes for 16 hours, the fitting this curve to a Gompertz model and extracting the peak growth rate from the model. In all cases the bacteria were grown in LB media, with carbenicillin added when the engineered plasmids were present. When compared to PBP8, the bacteria with harboring engineered plasmids grew significantly slower, and this difference was most pronounced in the 37°C condition, where the temperature sensitive promoter would be active and expressing the modified curli material. The error bars show the standard deviation of the samples (n>=12), and each plasmid bearing sample was compared to the PBP8 control with a two-sample t-test assuming unequal variances, * p<0.05, ** p<0.01, *** p<0.001. Assays performed by Anton Kan.

Supplementary methods: Synthetic plasmid retention in the mouse gut

7 female 8- to 9-week-old C57BL/6NCrl mice, obtained from Charles River Laboratories, were randomly split into 3 experimental cohorts: PBP8+pKAG, PBP8+pL6FO [-IPTG] and PBP8+pL6FO [+IPTG]. Housing and feed were the same as described in the main methods section. 3 days prior to initial administration of bacteria, the drinking water was supplemented with 0.5 g/L chloramphenicol, and all subsequent water contained chloramphenicol. For the +IPTG cohort, the water was supplemented with a further 10mM IPTG from day 0 onwards. Bacterial suspensions were prepared in advance by growing to mid-exponential phase (OD600 of 0.5) at 37°C (shaking at 225 RPM), pelleting the cells, resuspending to OD600 of 10 in PBS supplemented with 20% sucrose and 10% glycerol. All mice were gavaged with 10^{10} CFU of the relevant strain on day 0 of the experiment.

Fecal pellets were collected and weighed on day 1, 2, 3, 5, 7, and 9. Immediately following daily collection of fecal pellets, each sample was homogenized in 1 mL of PBS, serially diluted, and plated in quadruplicate to enumerate colony forming units (CFU). Samples were plated on two types of LB agar plates - 25 µg/mL chloramphenicol-only plates (LBC) and 25 µg/mL chloramphenicol + 50 µg/mL kanamycin plates (LBCK). While all PBP8-derived strains carried a chromosomal *camR* resistance gene, only the synthetic plasmid bearing population harbored the *kanR* resistance gene. Total PBP8 bacterial density was found by counting colonies on LBC plates and normalizing by the weight of fecal matter sampled. Plasmid retention rate was estimated by calculating the cell density from LBCK plates and dividing by the density from LBC plates.

Appendix C

Chapter 3 supplementary information

Supplementary Table C1: Strains and plasmids

(a) Strains		
Strain	**Description/comments**	**Reference**
Mach1	*E.coli*, str.K-12 F$^-$ ϕ80(*lacZ*) ΔM15 Δ*lac*X74 *hsd*R($r_K^- m_K^+$) Δ*recA*1398 *end*A1 *ton*A, cloning strain	Thermo Fisher Scientific
MC1061	*E. coli*, str. K-12 F$^-$ λ$^-$ Δ(ara-leu)7697 [araD139]B/r Δ(codB-lacI)3 galK16 galE15 e14$^-$ mcrA0 relA1 rpsL150(Str^R) spoT1 mcrB1 hsdR2($r^- m^+$)	Casadaban et al., 1980
PBP8	*E. coli* Nissle 1917, Δ*csgBACDEFG*::*Cm^R*	Praveschotinunt *et al.*, 2018
2457T	*Shigella flexneri* 2a strain 2457T	Mills et al., 1992
E2348/69	*E. coli* strain O127:H6/**EPEC** ; widely used as a model for EPEC infection	Iguchi et al., 2009
EDL933	*E. coli* strain O157:H7/**EHEC**, *stx1 stx2 eae espP/pssA hly$_{EHEC}$*)	O'Brien et al., 1983
E22	*E. coli* strain O103:K-H2/**REPEC**, *rha*-	Camguilhem et al., 1989
3014-2	*E. coli* strain O153:H-/**REHEC**, sorbitol+	Garcia et al., 2002
DBS770	*Citrobacter rodentium,* derivative of strain ICC168, chloramphenicol resistant. Lysogenized with Stx2dact-producing phage ϕ1720. Produces Shiga toxin.	Mallick et al., 2012
DBS771	*Citrobacter rodentium*, chloramphenical resistant, kanamycin resistant, *stx2dact*. DBS770 with KanR cassette inserted into prophage *stx* genes. Does not produce Shiga toxin.	Mallick et al., 2012
E10	*E. coli* strain O119:H6/EPEC wt	Giron et al., 2002

Supplementary Table C1 (Continued):

(b) Plasmids		
Plasmid	**Description/comments**	**Reference**
pL6FO	CsgA-VHH expression vector, contains synthetic curli operon *csgBACEFG* under an IPTG-inducible promoter.	Kan *et al.*, 2019
pGex-2t	Vector for EHEC intimin GST-Int465 (*eaeA*, EDL933 (C-terminal 465aa, bp1636-3082) (Yu and Kaper, 1992)); EPEC GST-TirM (E2348/69 TirM 120aa) (Perna *et al.*, 1998); REPEC GST-TirM (E22 TirM 120aa)	Pharmacia, Liu *et al.*, 1999
pDEST15	Vector for *C. rodentium* intimin GST-Int400C (*eaeA, Citrobacter rodentium* strain DBS100 (C-terminal 400aa));	Thermo Fisher Scientific
pGex-4t2	Vector for EPEC intimin (*eaeA*, E2348/69 (C-terminal 400aa)); REPEC intimin (*eaeA*, E22 (C-terminal 400aa))	Pharmacia
pET15b	Vector for EPEC His-TirM (E2348/69 TirM 120aa) (Perna *et al.*, 1998)	Novagen
pDest17	Vector for REPEC His-TirM (E22 TirM 120aa); *C. rodentium* His-TirM (DBS100 TirM 120aa)	Thermo Fisher Scientific
pMalc2	Vector for EHEC intimin MBP-Int395 (*eaeA*, EDL933 (C-terminal 395aa, bp1618-2082) (Liu *et al.*, 2002))	New England Biolabs, Liu *et al.*, 1999
pUC19	Vector for intimin expression in MC1061 (*eaeA*, full-length, EDL933 bp -30-3011; *eaeA*, full-length, JPN15 bp 1614-4537) (Yu and Kaper, 1992)	Yanisch-Perron *et al.*, 1985

Supplementary Table C2: Aligned sequences of VHHs binding pathogenic *E. coli* virulence factors. Complementarity determining regions (CDR1, 2 and 3) are highlighted and appear in order from left to right.

Anti-Fla:

JUV-B11: SGGGLAQPGGSLRLSCTSTGHT-LDDYAIGWFRQAPGKERERVACASASGI-TTNYADSVKGRFTISRDKAKNMVYLQMNSLQPEDTAVYYCAA---TPYYGDVCVRAAFESRGQGTQLTVSS

JUV-C4: TGGGLVQAGGSLTLSCVASGRA-VSSFAMGWFRQIPGREQRDFVAFIGDYGLTTYYANSVKGRFTISRSNAENTLYLQMNSLEFEDAAVYFCAA--RDAYRSGTTNPSAYDYWGQGTQVTVSS

JUV-E8: SGGGLVQAGGSLRLSCAASGRT-SSTYTMGWFRQAPGKEREFAAAIRSSGS-GTYYADSVKGRFTISRDGAKNTVYLQMNSLKPEDTAVYYCAA--RGNPIYSVYDVRTYDLWGQGTQVTVSS

JUV-G8: TGGGLVQPGGSLRLSCAASGSI-VSFNAMVWYREAPGKQREWVAQITPSSK--TMYKDSVKGRFTISSDNAKNMVYLQMNNLKPEDTAVYYCNGD----------RGVAWGPGTQVIVSS

JUV-H5: SGGGLVQPGGSLRLSCAASEMS-FSIKRMGWFREAPGKPREWVAQITPAGS--TNYAETVKGRFTISRDNAKNTVYLQMNSLKPDDTAVYYCNT----------LPGIAWGQGTQVIVSL

JWU-F3: SGGGLVQPGGSLRLTCVSSLSD-FRLTNMAWYRQTPGSERDVVAGISPNGI--TSYHASVQDRFNISRDNARKTLFLQMNSLKPEDSGVYYCNI----------RWGSLLPWGQGTQVTVSP

JWU-H4: TGGGSVQPGGSLRLSCAPSGFS-LAYYGVGWFRQAPGKEREALACISRFGD-DTYYADSAKGRFTVSRDNAKNTVYLEMNNLKPEDTGVYYCAAGWVVVTDESCGSGDAYNYWGRGTQVTVSS

JXE-B1: TGGGLVQAGDSLRLSCAASSGRNFSNYATGWFRQAPGKEQEFVASISRSGR-STYYADSAKGRFTISRDNARNTVYLQMNSLKPEDTADYYCAAHETQWPNGLGWVRGFYDLWGQGTQVTVSS

Anti-intimin:

JWS-H4: SGGGLVQAGGSLRLSCTTSASIFSGYRMGWFRQAPGKQREFVASIADG-QNTPYADSVMGRFTISRDNAKNTVYLQMSNLKPEDTAIYYCKS----------WGYTDWGQGTQVTVSS

JWT-C1: S-GGLVQPGGSLRLSCAASGFTLANSAIGWFRQAPGKGREAVSCISTSAGTTNYASSVKGRFTTARDNAKNMAYLQMDNLKSGDTGVYECAA------GWSIDCSGYILPAADVWGQGTQVTVSS

JWU-D8: TGGGLVQPGGSLRLSCAASGFYFSGYWMHWVRQVQGQGLKWVSGINIDDTKSSYTDSVKGRFTISRDNTKNTLYLQMDSLQPEDTGVYYCAR-------DRRAGQISGGYDPDYRGQGTQVTVSS

JWU-G8: SGGGLVQPGGSLRLSCATSGFYFAGYWMHWVRQVPGQGLEWVSGIDLGSTMLNYRDSVKGRFSISRDNAKNTVYLQMHSLKPEDTALYFCAR-------DRRAGATSGGYDPDVRGQGTQVTVSS

JXN-E2: T-GGLVPPGGSLQLSCTSSGFFLDYLGVAWFRQAPGNEREGVSCIDYTGDNIAYASSMKGRATISRDKDANTVTLEMNGLKPEDTAVYYCAAHRSATTYADGKYRCPLENEYDYWGQGIQVTVSS

Anti-Tir:

JVB-C6: TGGGLVQPGGSLTLSCAASGFSITENAMGWARQVPGKGLEWVSLV---------YSGGNTYYAESIEGRFTISRDNAKNTVYLRMTSLKPEDTGVYYCAA----REAVRVAGPPADVWGQGTQVTVSS

JVB-G4: SGGGLVEAGGSLRLSCSASGRASGDGHLGWFGDGALAWFRQAPGKEREYVAAVGHSGTDTYYSDSVKGRFTISRDNAKSMGYLQMDNLRPDDTGIYYCAL-----DLHLGQPGYDYWGQGTQVIVSS

JVB-G8: TGGGLVQAGGSLRLSCVASGFTFNDYVLTWFRQAPGKGREGVASI--------SPAFGNTFYADSVKGRFTITSDSAKKQVCLQMNNLKSEDTAVYFCAADPTVNFGPVLRAENYRHWGQGTQVTVSS

JVC-C6: SGGGLVQAGGSLRLSCAASGVTFVDYATGWFRQAPGKEREGIACI--------SNGASGSFVVNSVKGRFTTSTDNAKNTAYLQMNDLKPEDTATYYCAL-LGRTGGHCTDPNEYSYWGQGTQVTVSS

JVC-D10: T-GGLVQTGDSLTLSCVVSGRGFCDWAMAWFRQAPGKEREKVSAI--------GWSGQDTYYSEPAKGRFTISRDNAKNTVWLRMTNLKSEDTAVYYCAA-------ATRAYADYDYWGQGTQVTVSS

JVC-E5: TGGGTVQAGSSLRLSCTASGFTFTDAAMGWFRQAPGKEREFVAAI--------NWDSANKYYADSVKGRFTISRDNAKNTVYLEMTALKPEDTADYICAG--DAKLGHVATSDVWRFWGQGTQVTVSS

JVA-A1: S-GGSVEAGDSLRLSCAASGRGFGDGALAWFRQAPGKEREYVAAV--------GHSGTDTYYADSVKGRFTISRDNAKNMGYLQMDSLRPDDTGVYYCAL-----DLHLGQPGYDYWGQGTQVTVSS

JVA-C8: SGGGLVQAGGSLRLSCAASGFTFVDYGIAWVRQAPGKEREGVSCV--------STSNRSQWYADSVKGRFTISSDNAKNTVYLQMDNLKPEDTAVYYCTR----IDVGDNCRDGGGYRGQGTQVTISS

JVA-C9: TGGGLVQAGGSLRLSCVDSGRTFGDNAMGWFRQAPGKEREYVAAI--------GYNGDSTYYLNSVKGRFTISRDNAKNTLYLQMNSLQLEDTGVYICAA-------KTRVTTKEYDFWGQGTQVTVSS

JVA-D4: TGGGLAQAGGSLRLSCATSGFTFADYAIGWFRQAPGKEREGVACI--------STSDNGQYSADSVKGRFTISKDNAKKTVYLQMNSLKPDDTAIYYCHL--IDVGSNCRKGKDDGFWGQGTQVTVSS

JVA-F6: SGGGLVQAGDSLRLSCVTSGRSFSEAAMGWFRQAPGKERELMTSI--------GWNGDRTYYADSVKGRFTISRDNAKNTVYLQMNGLTPNDTAVYYCAA-------AVRASKGYEYWGQGTQVTVSS

JVA-D11: SGGGLVQAGGSLRLSCAASGFTFNDYAKAWFRRAPGKEREGISAI--------SSMGESTFYADSVKGRFTISSDSAKNTVYLQMNSLKPEDTAVYYCAADPTVKWGHVLREENYDYWGQGTQVTVVS

JVA-E10: SGGGLVQAGGSLRLSCVASGRTFSDYAMGWFRQAPGKEREFVAAI--------SWNGGSTYYADSVKGRFTISRDNAKNTMYLQMNSLKSEDTAVYYCIS-EGRGLRQVKTATDWEYWGQGTQVTVSP

JVA-G1: SGGGLAQAGGSLKLSCVASGFTFADYAIGWFRQAPGKEREGVACI--------SNSVGSTYYSDSVKGRVTISSDNVKKTVYLQMNSLKPEDTAFYYCAL--IDHGMNCRNSNGIHYWGKGTLVTVSS

Anti-EspA:

JXF-D7: SGGGLVQTGGSLRLSCSASGFALEYYAVGWFRQAPGKEREGVSCF--SGSDGSKFHAQFVKGRFTISLDKEKNTVDLTMNNLKPEDTAVYYCAVA-GPSDHQCDLGMTWYHRWGQGTQVTVSS

JYB-B1: TGGGLVQAGGSLRLSCAISGFSLGDYAIGWFRQAPEKEREGVAFS--GSLGNTY-YPDSMKGRFTISRDDAENAVYLEMKNLKPEDTAVYRCTS----GGTFDAVVLGLSTYWGQGTQVTVSS

JYB-B8: TGGGLVQPGGSLRLSCAASGFTFSTYWMYWVRQAPGKGLEWVSTI--DTGGSDTYYADSVKGRFTISRDNVKNTLYLQMDNLKPEDTALYYCGS---SRDIVIVTTLRDFDYWGQGTQVTVFS

JYB-D1: TGGGLVQAGGSLRLSCAASGNTFSYTAMAWFRQAPGKQRELVARI--SSGRGPTKYADSVKGRFTISRDNTLNTVWLQMNDLKPEDTAVYFCNT----LKYSGESSYIAGDSWGQGTQVTVSS

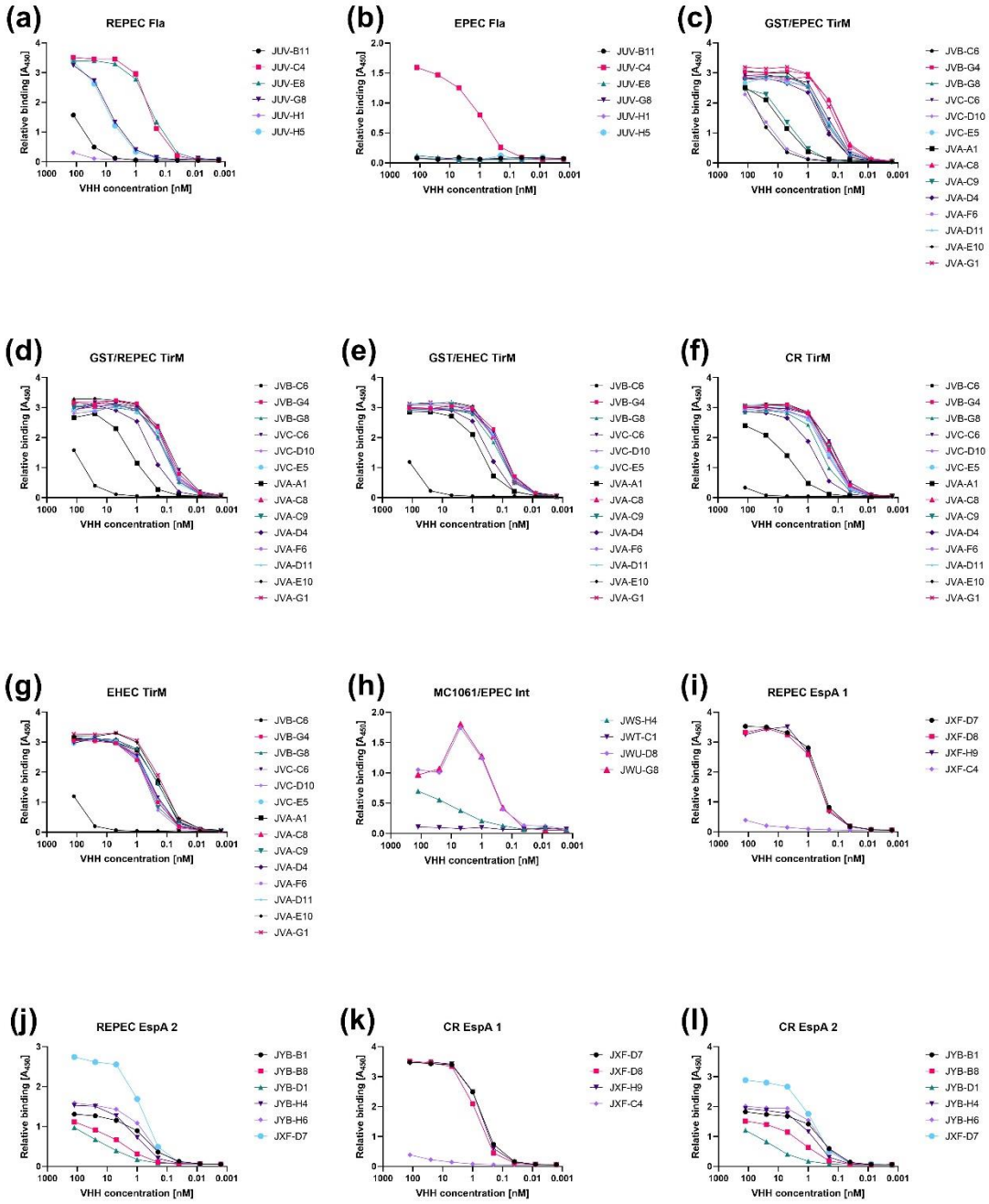
JYB-H4: SGGGLVQPGGSLRLSCAASGVTLEYYAIGWFRQVPGKEREGVSCI--STSGAGTNYADSVKGRFTISKDNAKNTVYLQMNSLKPEDTAVYYCAA-RDFTEVPIGGCEWEYDYWGQGTQVTVSS

JYB-H6: TGGGLVQPGGSLRLSCVASGFILDAYTIGWFRQAPGKEREGVASI--NGSGFSTNYADSVKGRFTISRDNAKNTVWLQMNSLKPEDTAVYYCAA--ALGLLTPRLESTPFDDWGQGTQVTVSS

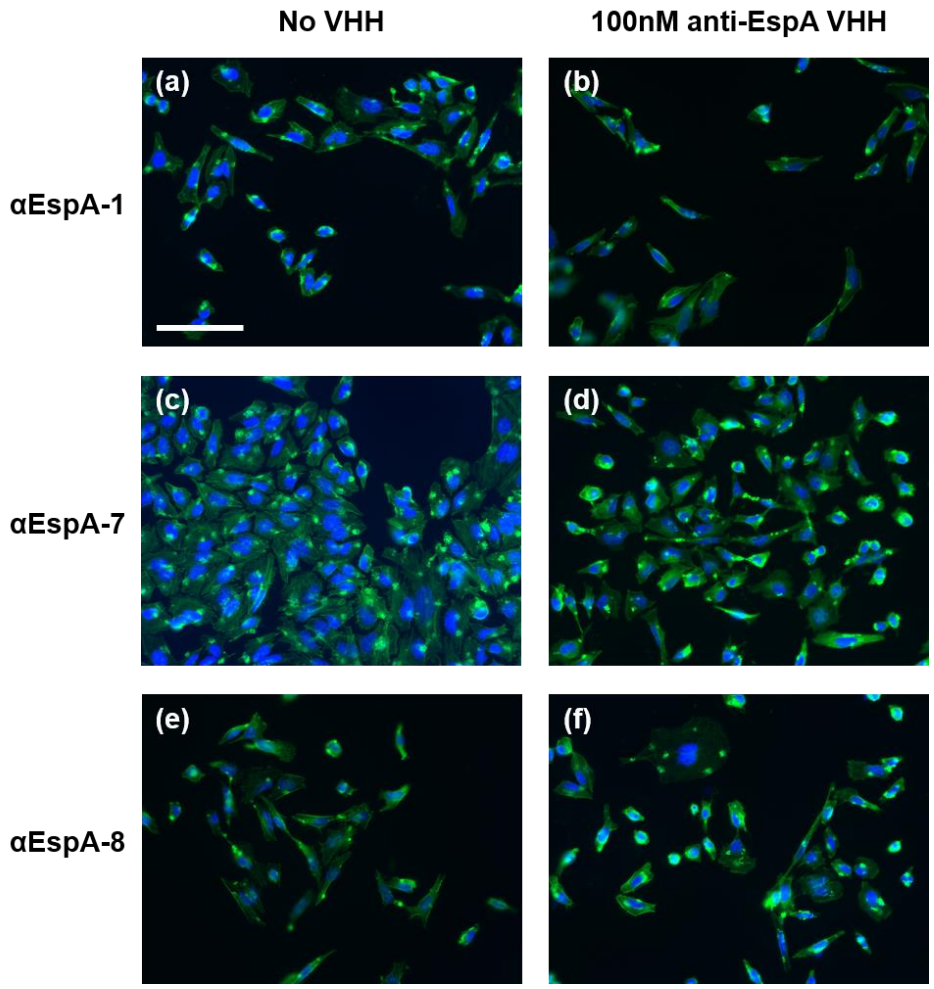
JYF-D8: SGGALVQPGGSLRLSCAASGINLDYYAIGWFRQAPGKERKGVSCI-SHVDDDRIYYSDSVKGRFTISRDNAKNTVYLQMNSLEPEDTAVYYCATA-GPSDYPCDLELWTYRRWGQGTQVTVSS

JXF-H9: SGGGLAQTGGSLRLSCAASGFRLEYYAVGWFRQAPGKEREGVSCL--SGSDGSTYNAEFAKGRFTISRDDAKNTVYLLMNSLKPEDTAVYYCAVA-GPSDYQCDLGTRWYHRWGQGTVVTVSA

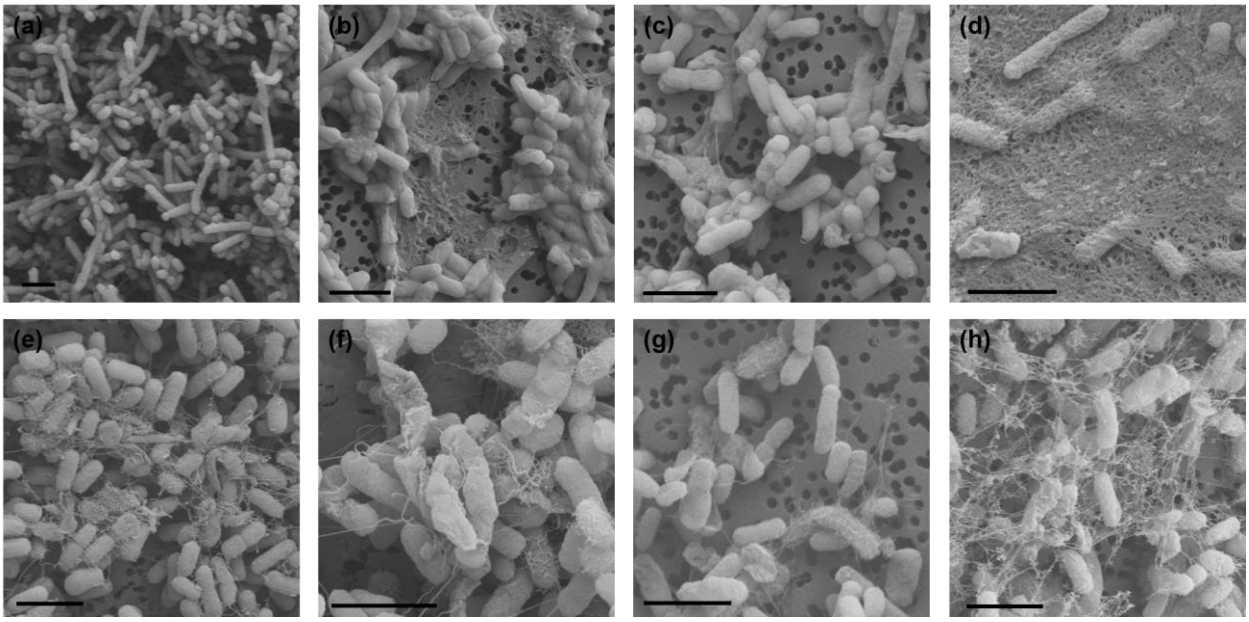
JXF-C4: SGGGLVQAGGSLRLSCVAYGRTANTFKEYAMGWFRHNPGEDHEFVGGISQNGDEAYFDDSVKGRFTPSRDNAKNTMYLQMNSLRPEDTAAYYCVAQRSSERLVGDMYSAMDSWGKGTLVTVST



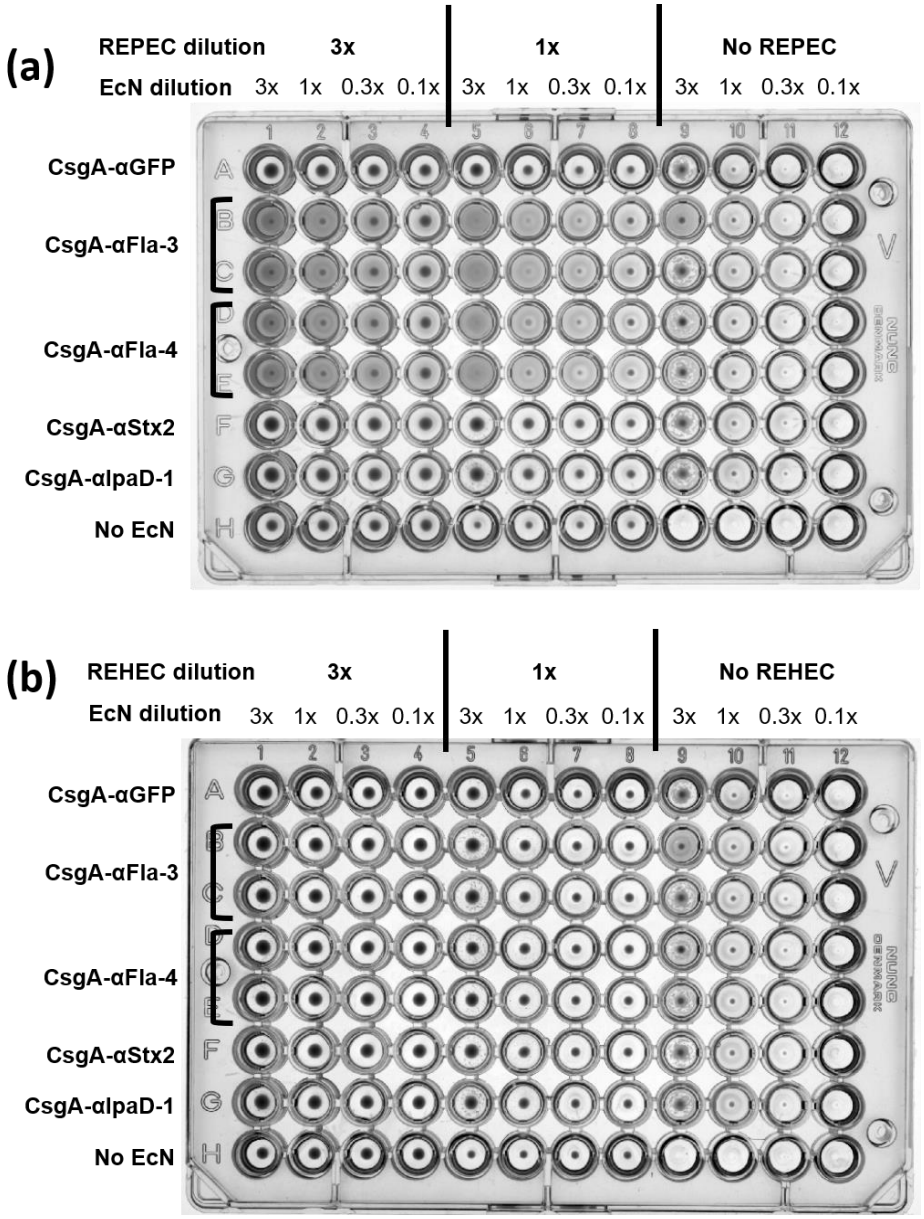
Supplementary Figure C1: Assessing the binding properties of selected VHHs via ELISA. Antigens used were homologues of Fla (a-b), Tir (c-g), Int (h) or EspA (i-l) corresponding to either EPEC (b, c, h), EHEC (e, g), REPEC (a, d, i, j) or *C. rodentium* (f, k, l). In each assay, antigen was either directly added to the plate in purified form (a-b, e-g, i-l), bound to the plate by an adsorbed noncompeting VHH (c-d), or displayed on the surface of MC1061 (h). Assays performed by Jacqueline Tremblay.



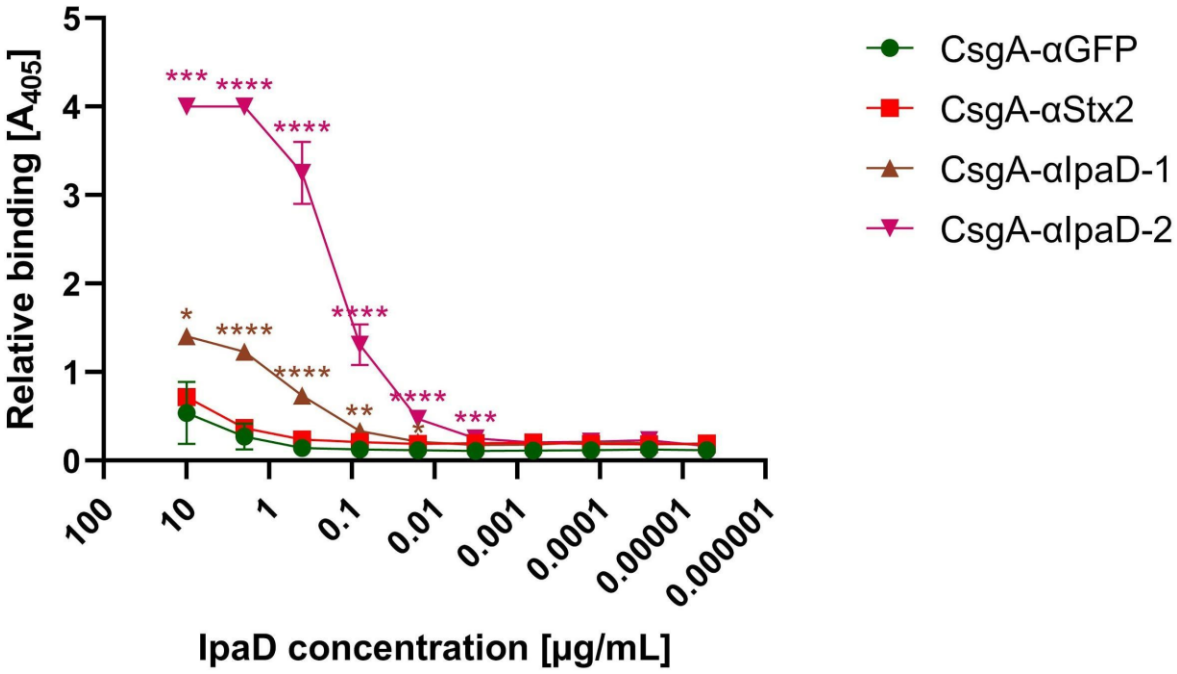
Supplementary Figure C2: Anti-EspA VHHs did not inhibit pedestal formation. HeLa cells were exposed to EPEC incubated with VHH, fixed and stained with DAPI (blue) and Alexa Fluor-488 Phalloidin (green). Similar to the "no VHH" negative control (a, c, e), all anti-EspA VHHs tested (b, d, f) resulted in the formation of pedestals (though not all anti EspA-VHHs were tested) (scale bar = 100 µm). Assays performed by Yousuf Aqeel.



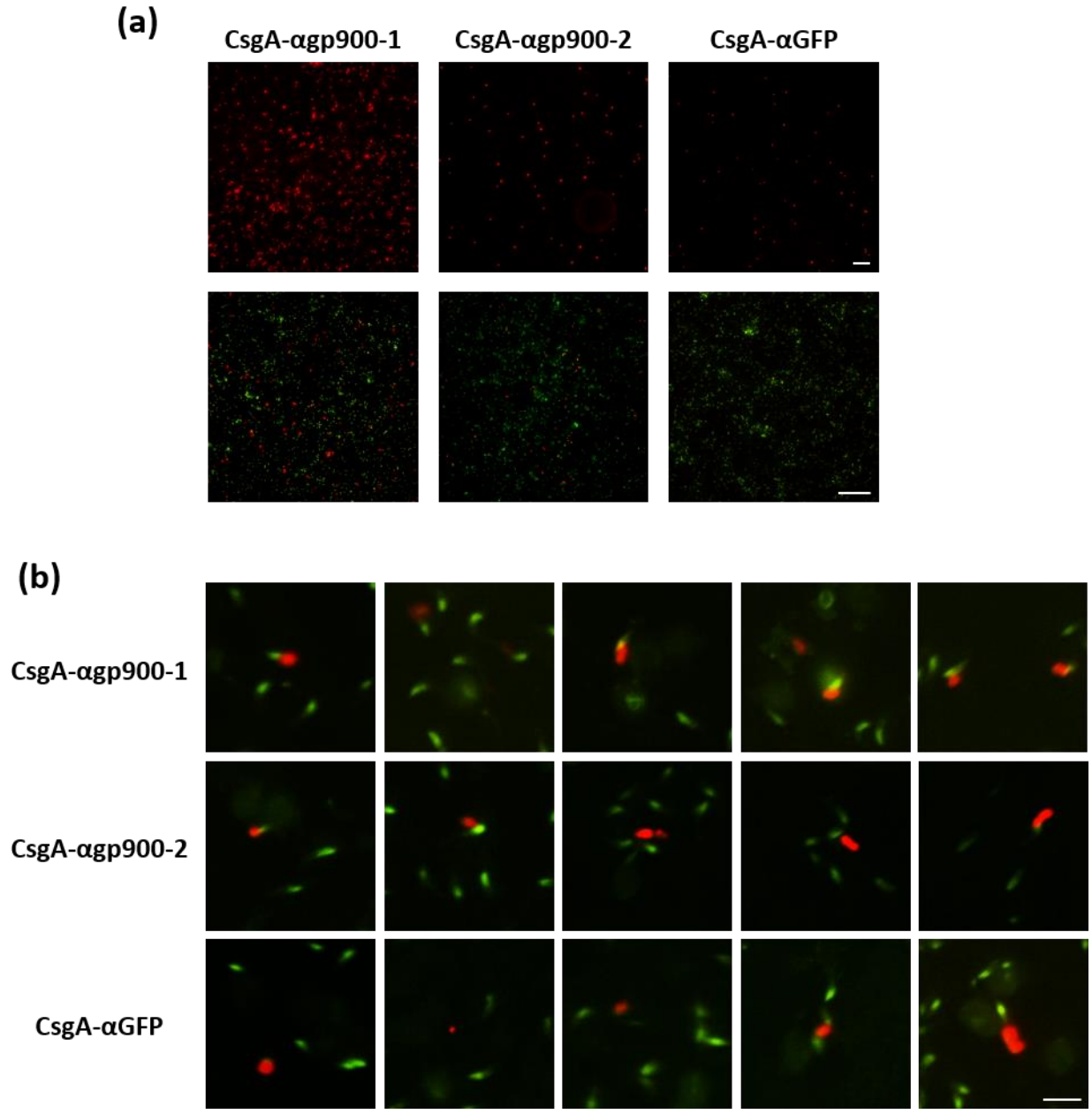
Supplementary Figure C3: Representative FESEM images of EcN expressing CsgA and CsgA-VHH. (a) EcN (PBP8) with no plasmid, expressing no curli fibers. (b-h) EcN expressing CsgA-VHH, exhibiting a range of fiber morphologies. (b) CsgA-αStx2, (c) CsgA-αInt-12, (d) CsgA-αInt-17, (e) CsgA-αFla-3, (f) CsgA-αFla-4, (g) CsgA-αIpaD-1, (h) CsgA-αgp900-2 (scale bar = 2 µm). Images captured by Avinash Manjula-Basavanna.



Supplementary Figure C4: EcN expressing CsgA-αFla can induce REPEC aggregation. Suspensions of EcN expressing CsgA-VHH were mixed with either REPEC (a) or REHEC (b) and allowed to settle overnight in conical 96-well plates. Aggregation was only observed when REPEC was mixed with CsgA-αFla-3 and -4. Assays performed by Jacqueline Tremblay.



Supplementary Figure C5: EcN expressing CsgA-αIpaD can bind soluble IpaD. ELISA demonstrated the ability of CsgA-αIpaD to bind soluble IpaD. EcN was adsorbed onto a well plate, followed by incubation with varying IpaD concentrations. Binding of IpaD to the adsorbed EcN was then detected by a specific non-competing VHH (JMK-H2, Barta *et al.* 2017), followed by an anti-Etag IgG-HRP conjugate. EcN expressing either CsgA-αIpaD-1 or CsgA-αIpaD-2 significantly outperformed the off-target negative control (CsgA-αGFP). Data presented as mean ± SD. Two-way ANOVA ($P < 0.0001$) was performed to test the presence of difference between conditions, P-values calculated by Welch's t-test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$. Assay performed by Justyna Jaskiewicz.



Supplementary Figure C6: EcN expressing CsgA-αgp900 exhibit increased attachment to and colocalization with *C. parvum* sporozoites. (a) Fluorescent micrographs demonstrating increased attachment of EcN (red) to *C. parvum* sporozoites, counterstained in green in the bottom panels (scale bar = 50 µm). (b) While EcN (red) expressing CsgA-αgp900-1 (and to a lesser extent CsgA-αgp900-2) consistently colocalize with sporozoites (green), the CsgA-αGFP negative control was often observed away from the green fluorescent foci, consistent with nonspecific binding (scale bar = 5 µm). Assays performed by Justyna Jaskiewicz.