Production, dynamics, and DNA binding capabilities of *Bacillus subtilis* natural competence associated pili

By

Jason Daniel Zuke

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Microbiology)

at the
UNIVERSITY OF WISCONSIN-MADISON
2024

Date of final oral examination: 11/30/2023

This dissertation is approved by the following members of the Final Oral Committee:
Briana M. Burton, Associate Professor, Bacteriology
Katrina T. Forest, Professor, Bacteriology
Jade Wang, Professor, Bacteriology
Daniel Amador-Noguez, Associate Professor, Bacteriology
Tu-Anh Huynh, Assistant Professor, Food Sciences

Contents

Acknowledgements ..... **iv**
Abstract ..... **vi**
Chapter 1: DNA transport to the cytoplasmic membrane during natural competence; an introduction **1**
- **Overview of bacterial gene transfer mechanisms** **2**
- **Mechanisms of bacterial horizontal gene transfer** **4**
- **Potential uses of DNA imported during natural competence** **7**
- **Natural competence directly affects society** **8**
- **Barriers to DNA entry during natural transformation** **9**
- **Development of the earliest natural transformation models** **11**
- **Identification of proteins mediating DNA translocation across the Gram-positive cell wall** **14**
- **Mechanics of pilus biogenesis and relationship to natural transformation** **17**
- **Evidence for pilus-DNA binding capabilities *in vitro* and *in vivo*** **19**
- **DNA binding/translocation by dynamic pili is confirmed via fluorescence microscopy** **21**
- ***Bacillus subtilis* competence pili, an enduring enigma** **23**

Chapter 2: Cysteine substitution mutagenesis of the *Bacillus subtilis* pilin homologue ComGC reveals residues amenable to fluorescent labeling with maleimide-conjugated dyes **24**
- **Abstract** **25**
- **Introduction** **26**
- **Results** **31**
- **Discussion** **36**
- **Materials and methods** **41**

Chapter 3: Visualizing dynamic competence pili and DNA capture throughout the long axis of *Bacillus subtilis* **48**
- **Abstract** **49**
- **Importance** **49**
- **Introduction** **50**
- **Discussion** **60**
- **Materials and methods** **68**
- **Acknowledgments** **80**

Chapter 4: Discussion **81**
- **Overall significance of this work** **82**
- **Natural competence pilus length control and retraction rates** **83**
- **DNA binding to natural competence pili** **85**
- **The necessity of the ComGC$_{Bs}$ disulfide bond** **86**
- **Future directions** **88**

Appendix A: Heterologous production of the ComGC soluble domain for solution NMR ..91
Introduction ..92
Results ..93
Discussion ..95
Materials and methods ..96
Acknowledgments ..98
Figures ..99
Figure 1 ..99
Figure 2 ..100
Figure 3 ..101
Figure 4 ..102
Figure 5 ..103
Figure 6 ..104
Figure 7 ..105
Figure 8 ..106
Figure 9 ..107
Figure 10 ..108
Figure 11 ..109
Figure 12 ..110
Figure 13 ..111
Figure 14 ..112
Figure 15 ..113
Tables ..114
Table 1 ..114
Supplementary material ..116
Supplementary figures ..116
Supplementary Figure 1 ..116
Supplementary Figure 2 ..117
Supplementary Figure 3 ..118
Supplementary Figure 4 ..119
Supplementary Figure 5 ..120
Supplementary Figure 6 ..121
Supplementary Figure 7 ..122
Supplementary Figure 8 ..123
Supplementary tables ..124
Supplementary Table 1 ..124
Supplementary Table 2 ..127
Supplementary Table 3 ..128

Supplementary Table 4 ..... **130**
Supplementary Table 5 ..... **131**
Supplementary Table 6 ..... **132**
Supplementary Table 7 ..... **133**
Supplementary Table 8 ..... **134**
Supplementary strain construction **135**
Chapter 2 ..... **135**
Chapter 3 ..... **143**
Appendix A **145**
Supplementary plasmid construction **146**
Chapter 3 ..... **146**
Appendix A **147**
References **148**

Acknowledgements

First and foremost, I want to thank my lovely wife Song for all the support she has given me over the last six years while I have engaged in sometimes life consuming endeavors. Even when we were split apart by pandemic travel restrictions prior to her immigrating to the US, Song would be there for me when my experiments weren't working or I was freaking out about a committee meeting, despite the fact that she was 13 or 14 hours ahead of me depending on the season. Her presence in my life has made me a better version of myself: a better husband, a better scientist, and a better man. She has been instrumental in helping me beat my insomnia, an ailment that had kept me from reaching my true potential for as long as I can remember, and her outgoing nature has encouraged me to finally figure out the sources of my chronic foot and back pain. When I think back to my health situation in my first year compared to now, I have made such progress in being able to lead a normal life again. I love you, Song!

Secondly, I need to shout out my parents Jeff and Mary Zuke who have always been supportive of my academic career and pursuit of knowledge. They instilled in me from day one a work ethic and drive that has allowed me to even be writing this document in the first place. Without their help, I would be a shell of my current self, so I couldn't have done this without them.

Next, I'd like to give a special thank you to my advisor Briana Burton, who has somehow managed to guide me through an absolutely insane number of problems in my thesis without ever being overbearing or a micromanager. She clearly believed in me, even during my worst moments, and allowed me to grow into the pretty darn good young scientist I am today. Her work ethic is an inspiration for me to not rest on my accomplishments and to keep driving towards a more successful future. It's obviously impossible to say, but I couldn't have seen myself being as successful in any other lab on campus.

Dovetailing from Briana, I'd like to next thank all the members of the Burton lab (past and present) who made the lab environment so lively. Jonathan Lombardino, thank you for helping me with the limited amount of bioinformatics I needed to wrestle with for this thesis, as well as introducing me to Inkscape for making figures I'm actually proud of. James Finn, thank you for distracting me in lab that one time and making me add undiluted primers to my PCR, that was very cool of you. I'm glad I knocked that DNA extraction tube out of your hand that other time and forced you to re-centrifuge it. Kasia Gromek, thank you for all the assistance over the years with basically every facet of the lab, from your *E. coli* freezing parties to your purification of ComGC for the NMR data included in Appendix A. Tanya Falbel, thank you for taking the time to read and comment on all of the written work I've sent you over the last six years, your experience brings a unique perspective I don't get from anyone else. Dom Smith, thank you for always being willing to make media/chemical solutions whenever we need it, that's super helpful. I hope you can stop needing to come into lab super late at some point! Heidy Elkhaligy, thank you for taking your personal time to make a ComGC pilus filament model for me, I really couldn't figure out how to do it. Mitch Armstrong, thank you for being such a goofy character in lab, it's honestly really nice to have someone like you in an otherwise serious/tense environment. I'm running out of steam here, so I apologize if I didn't mention you, and I'll just say thanks again!

Finally, I'd like to thank my committee for being really patient with me, even when my thesis was stalling out and it looked like I'd need to switch gears. You all gave me great feedback at my committee meetings, and you applied the pressure I needed to break out of my slump without being hostile. I made the right decision selecting each of you for my thesis committee.

Abstract

The transfer of genetic material between organisms is a fundamental aspect of life. Parent-to-offspring genetic transfer, or vertical gene transfer, is absolutely required for species propagation. Not all genetic transfers occur in this fashion, however. The acquisition of genetic elements from non-parental organisms, or horizontal gene transfer (HGT), is commonplace across the tree of life. Nowhere is this more evident than in the bacterial domain. Here, numerous HGT mechanisms operate to mediate genetic transfers from organism to organism, including phage transduction, conjugation, and natural transformation. While each mechanism is distinct from one another, they all share a common problem: Long, negatively charged, hydrophilic nucleic acid polymers cannot normally cross intact bacterial cell envelopes. Therefore, each of these HGT mechanisms has uniquely evolved to overcome this challenge.

The existence of genetically tractable, naturally transformable model organisms has enabled detailed study of the mechanisms involved in mediating DNA transit across the cell envelope during natural transformation. Natural transformation is the uptake of free DNA from the environment into the cytosol and the subsequent integration of this DNA into the recipient cell's genome, all of which is accomplished by specialized proteins produced by the recipient cell. The vast majority, if not outright all, of the proteins necessary for natural transformation have been identified in both Gram-negative and Gram-positive organisms. How these proteins operate mechanistically to transport DNA across bacterial cell envelopes has been an ongoing question for decades.

This thesis will focus on how DNA is transported across the Gram-positive cell wall, the first barrier encountered on the DNA's path to the cytosol, which is comprised of a dense meshwork of crosslinked peptidoglycan/teichoic acid. In the naturally transformable model organism

Bacillus subtilis, expression of only 10 genes is needed for DNA to cross the cell wall barrier in *B. subtilis*, with 80% of these genes encoding proteins homologous to essential pilus biogenesis proteins. Pili are thin, cytoplasmic-membrane-anchored filaments that span cell envelopes into the extracellular space. Naturally, it has long been hypothesized that *B. subtilis* (and other Gram-positives) employs a pilus system to breach the cell envelope and mediate DNA entry, but pili have never been observed on the surface of naturally transformable *B. subtilis*.

This thesis seeks to address this shortcoming while simultaneously exploring how these putative pili mediate DNA transport across the cell wall. Recently, a method was developed that allows for the fluorescent labelling of dynamic protein filaments in bacteria, the maleimide labeling method, which is compatible with live cell microscopy. Maleimide labeling involves cysteine substitution mutation of the protein subunits comprising the filament, on a residue that is solvent exposed. During filament production, a fluorescent maleimide conjugate is added to the medium, which will covalently link to the cysteine thiol on each monomer in the filament, making the structure fluorescent.

In chapter 2 of this thesis, we detail the production of a *comGC* (gene encoding the major structural protein of the putative pilus) cysteine substitution panel to search for labelable mutants. After screening Δ*comGC*^{WT} complementation strains expressing each of the 41 successfully made *comGC*^{Cys} variants, we found that 9 mutants maintained transformability within 500-fold of wildtype, providing a solid foundation for future maleimide labeling studies. In chapter 3 of this thesis, we demonstrate that strains expressing both *comGC*^{WT} and *comGC*^{E56C} or *comGC*^{S65C} maintained full transformability, improving on the original complementation strains. Both co-expression strains were capable of pilus production during natural competence, as evidenced by maleimide labeling and widefield epifluorescence microscopy, demonstrating

for the first time that *B. subtilis* does produce these structures. In experiments where labeled pili and fluorescently labeled DNA were imaged together, we find that these pilus structures could bind the DNA. During imaging experiments that tracked pili through time, we found that some pili had the ability to dynamically retract back towards the cell bodies. Together, these data suggest a model wherein *B. subtilis* produces pili during natural competence, the pili bind to environmental DNA, and then pilus retraction brings the DNA across the cell wall, through the pore created by the pilus extending originally.

Chapter 1: DNA transport to the cytoplasmic membrane during natural competence; an introduction

Jason D. Zuke wrote this chapter; Briana M. Burton, Jonathan Lombardino, and Tanya G. Falbel edited this chapter and gave helpful criticism

Overview of bacterial gene transfer mechanisms

Vertical gene transfer (VGT)

The transfer of genetic material from one bacterium to another is a fundamental aspect of the bacterial cell cycle. Gene transfer from parent to offspring, referred to as vertical gene transfer (VGT), is obviously necessary for the viability of any newly formed bacterium. During VGT, genetic mutations that arise in the parent cell will generally be passed on to the offspring, given that the offspring cell receives a copy of the parent genetic material containing those mutations (Fig. 1A). Natural selection can then act on these mutations, favoring the fixation of alleles that confer fitness advantages, and potentially novel beneficial traits, to recipient offspring (1).

The evolution of novel beneficial traits by mutation and subsequent VGT can be a slow, uncertain process. Take, for example, the fascinating evolution of an *E. coli* aerobic citrate utilization pathway found during the Lenski Long-Term Evolution Experiment (LTEE) (2, 3). A continually subcultured, clonal *E. coli* population grown in minimal glucose medium accumulated mutations sufficient for aerobic growth on citrate as the sole carbon source after ~ 30,000 generations (3). As of 2017 when the generation number hit the 65,000 mark, none of the other 11 lineages grown concurrently in the LTEE had evolved this capacity. But what if there was a way for a lineage to evolve such a metabolic pathway in as little as a single generation? Horizontal gene transfer (HGT) mechanisms are the key to such rapid evolutionary innovations.

Horizontal gene transfer (HGT)

Bacterial HGT, in contrast to VGT, is the transfer of genetic material between a cell and another cell that does not depend on a direct parent-offspring relationship (Fig. 1B). This enables the sharing of beneficial, evolved traits between multiple lineages. In practice, this means that traits

which were evolved in one lineage can both continue to be passed on to subsequent generations via VGT, and also passed from cells of one lineage to cells of another (lacking that evolved trait) via HGT. Bacteria can exploit multiple mechanisms of HGT, including phage transduction, conjugation, natural transformation, and so-called “non-canonical” mechanisms (4, 5).

Mechanisms of bacterial horizontal gene transfer

Phage transduction

The first HGT mechanism noted above, phage transduction, is dependent on bacteriophages (Fig. 2A). Phage transduction occurs when host DNA segments are inadvertently packaged into infectious virions following phage infection and translation of virion structural proteins (5, 6). These infectious particles can deliver their encapsulated bacterial DNA into other bacteria by the phage's standard entry mechanism (6). Once the virion's bacterial DNA has entered the new host cell, this transferred DNA can be integrated into the host genome by autonomous replication (in the case of plasmids) or insertion into the host chromosome by DNA repair pathways such as homologous recombination (6). Phage transduction has high utility in laboratories, where it is routinely employed for genetic engineering of organisms like *E. coli* when chromosomal segments need to be transferred between strains (7). Not all HGT mechanisms need phages to operate, however.

Conjugation

The next prominent HGT mechanism, conjugation, mediates the transfer of plasmids and other mobile genetic elements from one bacterial cell to another without the need for phages (Fig. 2B) (5, 8). In this system, dedicated protein machines replicate one strand of the DNA to be transferred (donor DNA) by the "rolling circle" replication mechanism (8). Rolling circle replication of the donor DNA is coupled to membrane fusion of the donor cell and a nearby recipient cell, which occurs by production of a type IV secretion system filament that forces donor and recipient cells into close contact with one another (8). The fused membranes create a channel through which donor DNA can flow into the recipient cell, where the cognate strand is synthesized for the stable integration of the donor DNA in the recipient cell (8). Prior to

discussing the most import HGT mechanism for this thesis, natural transformation, newly emerging mechanisms of HGT should be discussed.

Non-canonical HGT

A group of recently discovered HGT mechanisms, dubbed the "non-canonical" mechanisms, are an underappreciated part of the HGT landscape. Essentially, a mechanism of HGT is deemed non-canonical if it falls outside of the standard three mechanisms (namely phage transduction, conjugation, and natural transformation) (5). Two interesting examples are membrane vesicle and nanotube transfer, both involving apparent fusions of membranes to operate (Fig. 2C) (5). In the membrane vesicle mechanism, small assembles of membrane fragments (membrane vesicles) spontaneously form around environmental DNA fragments to create DNA-filled membrane particles (4, 5). These particles can fuse with the membranes of other cells, delivering the encased DNA into the outer cell periphery. For the nanotube transfer mechanism, cells grown on solid medium for extended time periods can form what appear to be membranous channels between each other when in close proximity (9). These channels have shown the capacity for delivery of both protein and DNA into adjacent cells (9).

Natural transformation

The final canonical HGT mechanism, natural transformation, is the overarching subject of this thesis (Fig. 2D). Prior to natural transformation occurring, bacterial cells will differentiate into a physiological state called "natural competence" or "genetic competence" (this thesis will henceforth refer to this cell state as natural competence) (10). During the transition to natural competence, cells will undergo a marked shift in gene expression, activating and repressing a wide range of genes (11). The end result of a shift to natural competence is the production of proteins within the naturally competent cell that mediate transformation, or the import of free

environmental DNA into the cytosol and the subsequent insertion of the imported DNA into the genome by homologous recombination (10). Natural transformation at the most basic level, therefore, is simply the transformation of naturally competent cells.

Potential uses of DNA imported during natural competence

DNA damage repair

There are three primary, potential benefits to importing DNA during natural competence (12). First, imported DNA could aid in DNA damage repair (13, 14). Imported DNA homologous to the chromosome was originally thought to mediate repair of damaged chromosomal segments by serving as templates for homologous recombination (12). However, transformation of a small *Haemophilus influenzae* chromosomal fragment improved survival of UV-irradiated cultures equally to transformation of purified genomic DNA, suggesting that transformed DNA must mediate DNA damage repair via an alternative mechanism (15).

Nutrient/energy source

In cases where DNA damage repair is not a priority, the imported DNA can be catabolized and used as a source of carbon, nitrogen, phosphorus, and energy (12, 16, 17). Alternatively, constituent nucleotides from the imported DNA can be recycled for DNA synthesis, thereby potentially saving cellular energy that would otherwise be used to produce nucleotides *de novo* (16).

Beneficial genetic element transformation

Finally, if the imported DNA is heterologous to the chromosome, novel genetic elements can be incorporated during integration (12). Such spreading of genes can significantly alter their frequency in populations, consequently contributing to bacterial evolution (5, 18). Although DNA import during natural competence is clearly important for bacteria, this process has many implications for human societies as well.

Natural competence directly affects society

Positive effects

Natural competence can have both positive and negative impacts on society. Numerous model organisms, both Gram-positive and Gram-negative, are genetically tracible in the laboratory due to their natural competence. *Bacillus subtilis* and *Streptococcus pneumoniae* strains are both produced via natural transformation, as are *Vibrio cholerae* and *Neisseria gonorrhoeae* strains (19–22). These organisms have been invaluable in elucidating the mechanisms underlying innumerable biological processes. Additionally, natural competence is vital for producing industrially relevant strains of species like *B. subtilis*. Due to the relative ease of genetically engineering *B. subtilis*, as well as the wealth of information available on the organism's physiology, many strains have been engineered to produce a wide variety of valuable biomolecules including B vitamins, high-value glycosaminoglycans like hyaluronic acid, and enzymes such as alpha-amylase (23–26).

Detrimental effects

On the other hand, gene transfer via natural competence can detrimentally affect human/animal health, as genes important for virulence, antibiotic resistance, and vaccine escape can be transferred between bacteria within a population. The majority of *Acinetobacter baumannii* clinical isolates tested are both naturally competent and resistant to commonly used antibiotics, and the transfer of carbapenem resistance between *A. baumannii* strains via natural competence has been experimentally validated (27, 28). Because of these detrimental effects, as well as the positive ones, it is crucial for the scientific community to have a thorough understanding of the DNA import process during natural competence. This starts with an understanding of the cellular barriers DNA faces during entry into the cytosol.

Barriers to DNA entry during natural transformation

The cytoplasmic membrane

Prior to any benefit being realized, bacteria must translocate DNA across the cell periphery. Such translocation is a monumental task for any species. This is readily apparent when the composition of the bacterial cell periphery is considered. Both Gram-positive and Gram-negative bacteria must mediate DNA crossing the cytoplasmic membrane during transformation. The cytoplasmic membrane is a lipid bilayer composed of roughly equal amounts of protein and phospholipid that maintains a net negative charge with high hydrophobicity (29, 30). DNA, a hydrophilic macromolecular polymer which bears negative charge along its entire length, faces numerous challenges in navigating the outer membrane, including insolubility, charge-repulsion, and steric hindrance. This is reinforced by the inability of *B. subtilis* protoplasts, cells lacking cell walls, to spontaneously transform exogenous DNA (31).

Gram-negative outer cell periphery

The permeability challenges faced by DNA at the outer cell periphery are quite different between Gram-negative and Gram-positive organisms. In Gram-negative species, the outer membrane first impedes DNA internalization. This lipid bilayer, composed primarily of densely packed phospholipids and lipopolysaccharide (LPS), presents an intensely hydrophobic and negatively charged surface (32, 33). DNA faces the same general issues getting through the outer membrane as it does the inner membrane, as described above. The cell wall, a thin (approx. 6 nm thick in *E. coli*) mono or dual-layered meshwork of glycan strands cross-linked together by short peptides, may provide an additional barrier to DNA entry for Gram-negatives (34, 35).

Gram-positive outer cell periphery

For Gram-positives, the cell wall first hinders DNA entry into the cell. Gram-positive cell walls are much thicker (approx. 35 nm thick for *B. subtilis*) than Gram-negative cell walls and are composed of 40 or more stacked peptidoglycan layers (36, 37). Unlike the Gram-negative cell wall, the glycan subunits of the Gram-positive cell wall are extensively modified with wall teichoic acids composed of repeating polyol-phosphate monomers (38). The phosphates within the teichoic acids impart a significant negative charge throughout and beyond the cell wall (38, 39). Imported DNA must therefore navigate tens of layers of Gram-positive cell wall, with each layer being extensively negatively charged. Like in the Gram-negative case, DNA faces both charge-repulsion and steric hindrance challenges when traversing the Gram-positive cell wall.

Development of the earliest natural transformation models

Discovery of natural transformation and early characterization efforts

Because of the implications natural transformation had for the evolution of virulence, and the opportunity this process inherently presented to study macromolecular transport against cellular barriers, many resources were devoted across the 20th century to better understand the mechanisms underlying DNA translocation during natural transformation. The study of natural transformation has a rich history, starting in 1928 when Frederick Griffith elegantly demonstrated that murine pneumococcal virulence could be transferred between strains (40). The discovery of antibiotic resistant and auxotrophic strains of numerous species provided convenient selectable markers for use in the earliest, detailed studies of natural competence and transformation from the mid-1950's to mid-1960's [*B. subtilis*: (41–44); *S. pneumoniae*: (45–47); *H. influenza*: (48–51); *N. meningitidis*: (52, 53)].

Antibiotic sensitive or auxotrophic strains could be transformed with DNA conferring antibiotic resistance or prototrophy, respectively, and then plated on selective medium to determine the number of cells transformed in an experiment. This strategy allowed for the rapid optimization of methods used for the transformation of the model organisms noted above by screening parameters such as which strain, growth medium/nutritional supplementation, growth temperature, growth phase, DNA concentration, and time of DNA exposure would result in maximum transformation. These studies were often performed concurrently with ^{32}P-labeled transforming DNA to track DNA uptake temporally, giving early insight into DNA uptake kinetics during transformation (43, 44, 47, 54, 55). DNA uptake in these examples is defined as irreversible DNA binding exhibited by naturally competent cells, wherein DNA becomes

resistant to exogenous DNase digestion. Uptake occurred rapidly (within 5 minutes) when competent cells were exposed to DNA and would mostly be completed within 15 minutes.

Temporal tracking of DNA forms generated during natural transformation

During the early 1960's to the mid 1970's, the application of density gradient ultracentrifugation techniques to radiolabeled donor DNA isolated during transformation led to the development of the overarching DNA uptake model still in use today [*B. subtilis*: (56–60); *S. pneumoniae*: (54, 61–63); *H. influenza*: (64, 65)]. In these experiments, naturally competent cultures were exposed to radiolabeled donor DNA for a short time to initiate DNA uptake. Further DNA uptake would then be terminated by DNase treatment, while still allowing the already-initiated uptake events to proceed. At varying time intervals after termination, DNA would be extracted from the cultures and subsequently subjected to chromatographic separation on various gradients and radiographic analysis. Extracted DNA species containing radioactive nucleotides from the donor DNA would separate within the chromatography columns differentially depending on multiple factors, including the species' molecular weight, form of DNA (single/double stranded, nucleotide/side), and association with the recipient chromosome (free/associated, covalently/reversibly bound). Crucially, this information could be used to quantitatively determine the fate of all donor DNA molecules through time after exposure of recipient cells to the donor DNA. This would grant a deep understanding of the sequence of events occurring during transformation to ultimately integrate donor DNA into the genome.

Together, the studies above suggested the following conserved mechanism of transformation: Shortly after exposure of naturally competent cells to exogenous DNA, DNA would breach the outer membrane and/or cell wall, then be fragmented and stably bound to the cell surface; once bound, one strand of the transforming DNA would be hydrolyzed while the other strand would

be transported through the cytoplasmic membrane into the cytoplasm (though which step came first was not resolvable); finally, homologous single-stranded DNA in the cytoplasm would pair with the recipient chromosome and eventually be covalently joined to the chromosome to complete the transformation process. Elucidating this sequence of events during transformation was a huge step forward in understanding the process, but an understanding of the proteins involved in transformation, and how they worked together mechanistically to achieve each step, remained unknown.

Identification of proteins mediating DNA translocation across the Gram-positive cell wall

First proteins discovered

Shortly after this transformation model was established, the first *S. pneumoniae* essential transformation protein was identified in 1975 (66, 67). This protein, an endonuclease responsible for the breakdown of the non-transforming DNA strand observed during DNA uptake, was identified by screening a mutagenized *S. pneumoniae* population for loss of external DNase capabilities (66, 68). Several mutants were isolated, each mapping to *S. pneumoniae*'s major endonuclease, *end* (66). Within the next decade, several transformation-deficient strains of *B. subtilis* would be identified, genetically mapped, and partially characterized (69–73). It would not be until the implementation of transposon mutagenesis that the bulk of transformation-associated proteins would be identified and characterized, including those likely responsible for translocation of DNA across the outer membrane and/or cell wall.

Transposon mutagenesis screen reveals a natural transformation pilus

In 1987, a transposon mutagenesis screen of *B. subtilis* for transformation deficient mutants was performed that not only identified many novel genes associated with transformation, but also gave early insight into their potential functions (74). One group of mutants mapping to the same locus were completely non-transformable and unable to bind exogenous DNA during natural competence. Further characterization of this locus revealed an operon (*comG*) composed of seven genes, each homologous to essential genes in Gram-negative type II protein secretion systems (T2SS) and type IV pilus (T4P) systems (Fig. 3A) (75–77).

The homology of ComG ORF 3 (now known as ComGC) with major structural pilin proteins from a diverse set of T4P and T2SS systems suggested that ComGC may oligomerize into a surface-exposed, pilus-like structure capable of interacting with environmental DNA (Fig. 3B) (76). Instead of mediating protein secretion through the cell wall and outer membrane as seen in homologous T2SS, the ComG proteins could be mediating DNA transport from outside of the cell through the cell wall (76). This idea was further supported by the presence of another essential competence protein (now known as ComC) in *B. subtilis* that is homologous to pre-pilin proteases essential for pilus biogenesis in other systems (74, 76). ComGC proteolytic processing was, indeed, later found to be dependent on ComC (78).

Characterization of the putative ComG pilus

Despite the strong bioinformatic and biochemical data supporting the production of ComGC pili during *B. subtilis* natural competence, these putative structures were not observable via imaging techniques such as electron microscopy, so their existence remained in question (80). To address this discrepancy, a 2006 investigation sought to biochemically characterize cell-envelope associated protein complexes containing ComGC that were produced during natural competence (80). Multiple ComGC-containing species were discovered within the cell wall fraction of competent cells upon Western blotting for ComGC following non-reducing SDS-PAGE. The apparent molecular weight of these proteins ranged from ~10 kDa (corresponding to ComGC monomer) up to the resolution limit of the gel, increasing by ~10 kDa along the entire lane. These data indicated that competent cells produce disulfide linked ComGC oligomers that extend outward past the cell membrane, exactly what one would expect if ComGC oligomerizes into pili.

With strong evidence for ComGC pilus biogenesis in hand, this same study assessed the genetic factors responsible for pilus biogenesis. Expression of *comG* and *comC* alone, in a strain background incapable of producing other natural-competence-specific proteins, was sufficient for production of ComGC oligomers. Expression of the genes encoding the dithiol oxidoreductase pair *bdbDC* was also required for ComGC oligomerization, though basal *bdbDC* expression (e.g. expression during planktonic growth) was sufficient (80). BdbDC are necessary for robust ComGC stability via formation of an intramolecular disulfide bond and are presumably instrumental in the formation of the intermolecular disulfide bonds found in the ComGC oligomer (80). The ability to induce ComGC pilus production in otherwise non-competent cells then allowed for a controlled investigation of the specific role of these pili during natural transformation.

Irreversible DNA binding to cells was achieved by expression of *comEA*, which encodes the integral membrane protein responsible for DNA binding at the cell membrane, alongside *comG*/*C* (81). Because ComEA localizes to the cell membrane and exerts its activity there, this heavily implied that the function of the ComGC pilus was to mediate translocation of DNA across the normally impermeable cell wall (81, 82). *comG* is dispensable for DNA binding to naturally competent protoplasts, i.e. cells lacking cell walls, lending additional support for ComG facilitating DNA translocation across the cell wall (82). The mechanisms underlying this translocation step, however, remained unclear.

Mechanics of pilus biogenesis and relationship to natural transformation

Production of pili

To fully appreciate how pili or pili-like structures could mediate DNA translocation across the outer cell periphery, the mechanics of pilus biogenesis must be established (reviewed in (83, 84)). A class of proteins, the prepilins, are produced and inserted into the cytoplasmic membrane. These 10-25 kDa proteins are embedded in the membrane by a conserved N-terminal alpha helix, which is flanked by a variable C-terminal region. Dedicated proteases cleave the prepilins at their extreme N-termini to form mature pilins. A cytosolic ATPase and a polytopic membrane protein then work together to couple ATP hydrolysis to the extrusion of pilins into a 6-9 nm wide, multiple μm long helical filament (the pilus) held together by protein-protein interactions between pilin monomers. The conserved N-terminal alpha helices are buried in the core of the filament, while the variable C-terminal domains are solvent-exposed along the length of the pilus. This cytoplasmic membrane-anchored pilus extends across the cell wall and outer membrane, through a secretin channel, into the extracellular space where the pilin C-terminal domains can interact with the environment. The pilus can be retracted back into the cytoplasmic membrane if a retraction ATPase, or a single bifunctional ATPase, is present in the cytoplasm (85). The homologous T2SS likely assemble pili in a similar fashion during protein secretion, although these structures do not extend beyond the cell envelope, leading to their designation as "pseudopili" (reviewed in (86)).

Hypotheses regarding pilus-mediated DNA import during natural competence

This understanding of pilus biogenesis led to two major hypotheses regarding how (pseudo)pili structures could translocate DNA into the periplasm. The first hypothesis viewed the structures as passively assisting in DNA transport, where their expression either provides a direct protein

channel through the cell envelope which DNA uses to enters the periplasm, or cycles of assembly and disassembly leave open paths in the cell wall/outer membrane that DNA can diffuse through (12). The second hypothesis saw the structures taking on a more active role where, after assembly, pili directly bind to DNA and guide the DNA across the cell envelope, most likely via retraction (12, 87).

Evidence for pilus-DNA binding capabilities *in vitro* and *in vivo*

In vitro evidence

By the early 2000s, a correlation between T4P production and natural competence had been described for numerous Gram-negative bacteria, including *Neisseria* and *Pseudomonas* species (88, 89). In these same species, the competence defect in pilus biogenesis mutants appeared to arise from an inability to bind extracellular DNA, as was observed for *B. subtilis comG* transposon insertion mutants (74, 89, 90). From the mid 2000's to early 2010's, further studies into Gram-negative T4P demonstrated unambiguous binding of DNA to T4P. Purified, intact *P. aeruginosa* T4P (noted above) were shown to bind both prokaryotic and eukaryotic DNA sequence non-specifically *in vitro*, and this interaction was disrupted significantly by preincubation of the T4P with a monoclonal antibody targeting the pilus tip, suggesting that the pilus tip binds DNA (91). Interestingly, purified pilin was unable to bind DNA, which pointed to DNA binding being an emergent property of the pilus quaternary structure. *N. meningitidis* T4P have also been found to bind DNA *in vitro* in the context of intact pili (92). In contrast to the *P. aeruginosa* T4P, the DNA binding of *N. meningitidis* T4P is greatly enhanced by the presence of a specific 12 bp sequence known as the DNA Uptake Sequence (DUS), and this specific binding is achieved by the presence of a low-abundance pilin, ComP, in the pilus fiber. Purified ComP contained intrinsic DNA binding capabilities, seen in electrophoretic mobility shift assays, which suggested that the pilus quaternary structure may not always be crucial for DNA binding.

In vivo evidence

A 2013 investigation into the *comG* operon of *S. pneumoniae* demonstrated, for the first time, direct *in vivo* evidence of a competence-specific pilus binding to extracellular DNA. The *S. pneumoniae comG* operon encodes seven genes, each of which shares extensive homology with those in the *B. subtilis comG* operon noted above (93). Both immunofluorescence and electron microscopy revealed that thin, extended structures were produced at the surface of competent cells, and that these structures were comprised at least in part of ComGC. The flexibility, length, and width of these structures was consistent with pili from T4P systems. Mass spectrometry analysis of affinity-purified pili confirmed that these pili were composed primarily of ComGC. Co-incubation of competent cells with bacteriophage lambda DNA, followed by electron microscopy, showed pili on the surface of *S. pneumoniae* cells in close association with the DNA both at the pili tips and along the fibers. These data together led to the conclusions that a dedicated type IV competence pilus composed primarily of ComGC is produced by *S. pneumoniae*, and that this competence pilus directly binds extracellular DNA to begin the internalization process. The latter conclusion is complicated by the potential for the cell fixation process, which is necessary for electron microscopy, to artifactually place T4P and DNA in close proximity. How DNA would be translocated into the periplasm once bound to competence pili, if binding was occurring, remained uncertain.

DNA binding/translocation by dynamic pili is confirmed via fluorescence microscopy

Maleimide labeling of the Vibrio cholerae (Gram-negative) natural competence pilus

The next breakthrough in this area came in 2017 from the development of a method to label dynamic, transformation-specific T4P from live cells (94, 95). This method works by first introducing a cysteine substitution into the main pilin protein, at a solvent-accessible residue to provide the pilin with a free thiol group. Upon production of this cysteine-substituted pilin, a maleimide-fluorophore conjugate is added to the growth medium. Maleimide is a thiol-reactive compound capable of forming covalent bonds with the free thiols found on each pilin monomer. The maleimide-fluorophore conjugate reacts with the available pool of pilin monomers, thereby labeling the pilin with whichever fluorophore is conjugated to the maleimide. The active assembly and disassembly of T4P comprised of these labeled pilins can then be observed in real time using widefield fluorescence microscopy.

This technique was successfully applied to the T4P of the Gram-negative *Vibrio cholerae* by introducing a cysteine into the major pilin PilA to make $PilA^{Cys}$ (96). Upon activation of natural competence, and subsequent addition of a maleimide dye to the growth medium, dynamic extension (up to 2.5 μm in length) and retraction of fluorescent T4P filaments were observed via widefield fluorescence microscopy (Fig. 4A). When a fluorescently labeled PCR product was added to the medium, these T4P filaments co-localized DNA at their tips (Fig. 5A), with pilus retraction bringing the DNA in close proximity to the cell periphery (Fig. 6A). Rates of natural transformation in *V. cholerae* mutants that produced DNA-binding-deficient T4P were severely reduced, indicating that DNA binding was crucial for transformation. Physical obstruction of T4P retraction, achieved by linking maleimide-biotin to the $PilA^{Cys}$ monomers and adding

neutravidin to the medium to form bulky T4P-neutravidin complexes, resulted in a similar drop in transformation rate, indicating that T4P retraction was also critical for transformation.

Maleimide labeling of the Streptococcus pneumoniae (Gram-positive) natural competence pilus

The same general strategy has been successfully employed for the Gram-positive organism *S. pneumoniae* with a *comG* operon homologous to *B. subtilis comG*, as mentioned above (93). *S. pneumoniae* also produced dynamic T4P filaments (Fig. 4B) that were capable of DNA binding (Fig. 5B), with retraction being similarly crucial for transformation (Fig. 6B) (97). Critically, regardless of any observed differences between these pilus systems, all available evidence points to a conserved mechanism of DNA translocation across the outer layers of the cell periphery during natural competence: Naturally competent cells produce pili that can access the extracellular space; these pili bind to DNA in the extracellular space; pilus retraction brings the bound DNA across the outer membrane and/or the cell wall into association with proteins essential for the next steps in uptake (Fig. 7).

Bacillus subtilis competence pili, an enduring enigma

But is this mechanism truly conserved, or is there more diversity waiting to be discovered? Currently, the accepted mechanistic model of DNA transport to the cytoplasmic membrane is based on one report of a single Gram-negative species (*V. cholerae*) and a single Gram-positive species (*S. pneumoniae*) (96, 97). Despite the fact that *S. pneumoniae* and *B. subtilis* share a homologous *comG* operon, for example, there are significant differences between oligomers of each species. Specifically, the essential disulfide bond found both intramolecularly in $ComGC_{Bs}$ monomers and the apparent intermolecular disulfide bonds present between $ComGC_{Bs}$ monomers within oligomers (80, 93). The inability to readily observe $ComGC_{Bs}$ pili from competent *B. subtilis*, while $ComGC_{Sp}$ pilus formation was readily observed in *S. pneumoniae* by both electron/ immunofluorescence microscopy, suggests that the structure of ComGC pili may vary significantly within the naturally competent Gram-positives. The subsequent chapters of this thesis will detail the thorough investigation of the *B. subtilis* ComGC pilus to begin addressing these questions.

Chapter 2: Cysteine substitution mutagenesis of the *Bacillus subtilis* pilin homologue ComGC reveals residues amenable to fluorescent labeling with maleimide-conjugated dyes

Jason D. Zuke designed and carried out experiments, analyzed data, and wrote this chapter.

Briana M. Burton assisted with conceptualization for this chapter

Abstract

Natural competence is the conserved ability for bacterial cells to produce a multiprotein machine that imports free, extracellular DNA into the cytosol. Gram-positive cells must translocate DNA across the peptidoglycan cell wall during this process. In the Gram-positive model organism *Bacillus subtilis*, multiple lines of evidence strongly suggest that *comG* and *comC* expression results in production of a filamentous, cell-envelope spanning, ComGC-based pilus complex which mediates DNA transport across the cell wall. However, evidence of *bona fide* pilus production in *B. subtilis* is lacking, and little is consequently known about the mechanism underlying ComG/C mediated DNA translocation. In this report, we leverage the protein structural modeling software AlphaFold2 to demonstrate that ComGC monomers probably adopt a traditional pilin tertiary structure. We detail the production of 41 *comGC* cysteine codon substitution mutants (*comGC*^{Cys}) that, based on the AlphaFold2 data, are most likely to produce functional ComGC pili amenable to fluorescent labeling with maleimide-conjugated dyes. Finally, we validate our *comGC*^{Cys} mutant panel and find 9 *comGC*^{Cys} mutants maintained transformability within 500-fold of wildtype.

Introduction

Across the bacterial domain, many species can import extracellular double stranded DNA into the cytosol via a specialized multiprotein machine (98). Cells that exhibit this DNA import ability are referred to as “naturally competent”. DNA imported by naturally competent cells can be used for a wide range of cellular needs (10). First, imported DNA can be metabolized to assist in meeting cellular nutritional and/or energetic requirements (16). *De novo* nucleotide biosynthesis is an energetically costly process, for example, so any nucleotides supplied by internalized DNA catabolism may conserve the cellular energy pool by obviating the need for additional nucleotides (16). Secondly, imported DNA may be used to stimulate DNA repair pathways (14). The exact mechanisms behind this phenomenon are unclear, but cell viability of UV-irradiated naturally competent *Streptococcus pneumoniae* cultures increases significantly upon addition of heterologous PCR product to the cultures post-irradiation (14). Finally, imported DNA may be integrated into the recipient cell’s genome via homologous recombination, increasing genetic diversity and potentially granting novel metabolic activities to the cell in a single evolutionary step (98). Such an event is called “natural transformation”. This is exemplified by the inter-strain transfer of antibiotic resistance traits in *Acinetobacter baumannii* via natural transformation (27, 28).

Whatever the ultimate fate of imported DNA may be, Gram-positive cells must all transport DNA through the cell wall during internalization. The Gram-positive cell wall presents a significant challenge for DNA transport. The effective pore size of the peptidoglycan meshwork comprising the cell wall is almost exactly as wide as standard B-form DNA, meaning that steric hindrance likely impedes DNA from freely crossing the cell wall (99). Another barrier for DNA entry is the maintenance of a net negative charge on the cell wall through wall teichoic acid

(WTA) modifications (38, 100). WTA, comprised of repeating polyol-phosphate subunits, is covalently linked to upwards of 10% of all peptidoglycan monomers within the cell wall, with the phosphates imparting significant negative charge (100). Extracellular DNA, also bearing negative charge along the entire polymer, faces charge-repulsion at the cell wall. Because the Gram-positive cell wall is composed of 40+ layers of peptidoglycan linked to WTA, DNA faces an impenetrable maze of small, like-charged pores on its way to the cell membrane (36, 37).

Despite this seemingly insurmountable problem, naturally competent cells can translocate DNA across the cell wall. In *B. subtilis*, protein products of the *comG* operon and *comC* gene mediate DNA transport across the cell wall barrier. *comG's* essentiality in this process is illustrated in DNA binding experiments to naturally competent cells lacking cell walls, i.e. naturally competent protoplasts (82). DNA binding to competent cells is abolished in *comG* mutants, however *comG* mutant protoplasts are capable of binding DNA at near-wildtype levels (82). This implies that *comG* expression is necessary for DNA breaching the cell wall. Additional experiments have shown expression of *comG*/*comC*/*comEA*, without any other natural competence associated genes, is sufficient for DNA binding to otherwise non-competent cells (though a constitutively expressed oxidoreductase gene pair *bdbDC* are also necessary) (80). ComEA, a known integral membrane protein responsible for binding DNA through its C-terminal domain at the cell membrane, cannot exert an effect on extracellular DNA (81, 82). BdbDC, while similarly membrane-localized, are necessary for ComGC stability and DNA binding (80, 101). Therefore, the ComG/ComC/BdbDC proteins appear to be both necessary and sufficient for translocation of extracellular DNA across the cell wall.

Bioinformatic analysis of the *comG* operon and *comC* gene suggested a potential DNA transport mechanism across the cell wall. Each gene within *comG/C* exhibits striking homology to the

essential components of both type IV pili (T4P) and type II secretion system (T2SS) pseudopili production systems (76). Both systems produce oligomeric protein filaments (pili or pseudopili) composed of helical assemblies of small pilin proteins which are anchored in the cytoplasmic membrane and span at least the length of the cell envelope (83, 84, 86, 102, 103). T4P systems produce pili that routinely extend multiple micrometers beyond the cell envelope into the extracellular space, whereas T2SS produce short filaments that typically do not extend past the outer cell envelope, granting these filaments the "pseudopili" moniker. Critically, T4P are dynamic structures that extend/retract via the insertion/removal of pilin monomers at the base of the filament (83, 84). Extension of pseudopili is required for secretion of T2SS substrates, while both the extension and retraction of T4P are often necessary for T4P function, as in the case of T4P-mediated twitching motility (83, 103).

Given the functions and mechanisms of action of T2SS pseudopili and T4P, the homology of *comG/C* to essential genes encoding components of these systems, and the known role of ComG/C in mediating DNA transport across the cell wall, ComG/C almost certainly work together to produce a pilus-like filament that creates a channel through the cell wall and somehow mediates DNA movement through that channel (76, 82). This is further corroborated by the biochemical characterization of oligomeric ComGC (pilin homologue) complexes found within cell wall fractions of competent cells, which are apparently stabilized by unusual inter-subunit disulfide bonds (80). How, precisely, these ComGC assemblies mediate DNA transport has been difficult to determine mechanistically, owing mainly to a lack of structural and microscopic data identifying and characterizing the structures (80).

In another naturally competent Gram-positive model, *Streptococcus pneumoniae,* the picture has become clearer. *S. pneumoniae* encodes homologous ComG and ComC proteins which are

produced during natural competence (93, 97). Multiple studies have found *bona fide* pili composed primarily of ComGC being produced by naturally competent cells, that not only are dynamically produced and retracted, but also bind extracellular DNA (93, 97). Pilus dynamism was established by the recently developed maleimide labeling method (95). Here, a cysteine codon substitution mutation is introduced into the gene encoding the main pilin subunit comprising the pilus (in this case *comGC*) and maleimide-conjugated fluorophores are introduced into the medium during natural competence. If the mutated ComGC residue is solvent-accessible, the maleimide will covalently react with the cysteine thiol, labeling all ComGC subunits. Dynamic pilus extension and retraction can then be visualized via widefield epifluorescence microscopy.

This has led to creation of a Gram-positive DNA translocation model wherein ComG/C produce pilus filaments using ComGC monomers, DNA binds to extracellular pili during pilus extension phases, and the DNA is subsequently pulled through the cell wall channel during pilus retraction phases (97). This model is currently based only on *S. pneumoniae*, so the conservation of this specific DNA translocation strategy across the Gram-positives is uncertain. Additionally, not only is there currently no direct evidence of extended pilus production in naturally competent *B. subtilis*, but the intermolecular disulfide bonds which apparently stabilize $ComGC_{Bs}$ oligomers are absent in *S. pneumoniae* competence pili, suggesting that the biogenesis and overall quaternary structure of $ComGC_{Bs}$ oligomers may differ significantly from $ComGC_{Sp}$ pili. To better understand if/how *B. subtilis* produces natural competence pili and the mechanisms underlying DNA translocation across the cell wall, we have adapted the maleimide labeling method for use with $ComGC_{Bs}$. Here, we use the protein structural modeling software AlphaFold2 to show that ComGC likely folds into a structure homologous to other,

experimentally determined pilus structures. Using this structural model as a guide, we generate 41 *comGC* cysteine codon substitution mutants (*comGC*^{Cys}) that are most likely to produce functional, maleimide-labelable ComGC pili. Finally, we screen our *comGC*^{Cys} mutants for complementation of Δ*comGC*^{WT} non-transformability and find 9 *comGC*^{Cys} mutants with transformability within 500-fold of wildtype that will serve as candidates in future maleimide labeling studies.

Results

ComGC likely adopts a typical pilin domain architecture

In an effort to rationally target *comGC* cysteine codon mutagenesis for downstream maleimide labeling, we sought to determine which ComGC residues would be most amenable to interaction with maleimide-conjugated fluorophores. The two most vital characteristics of these residues are that they must be solvent-exposed in the context of ComGC oligomers to promote maleimide accessibility, and these residues must not be essential for the oligomerization of ComGC monomers so that mutation does not disrupt oligomer formation. AlphaFold, a protein structural modeling system which incorporates machine learning to predict protein structure based solely on an input amino acid sequence, has given us invaluable insight into the structure of monomeric ComGC despite the current lack of experimental structural data (104).

Utilizing the Google ColabFold v1.5.3 webserver (running AlphaFold2), we have predicted the three-dimensional structure of monomeric ComGC (Fig. 8A). The structure is comprised of an extended N-terminal alpha helix depicted in red, flanked by a C-terminal domain depicted in blue, with a disulfide bond joining the two domains. Our confidence in this structure was greatly increased by the prediction of a disulfide bond, which matches experimental biochemical data indicating the presence of an intramolecular disulfide bond within ComGC monomers (105). Critically, this ComGC model is structurally homologous to experimentally derived pilin models from T4P of both *Pseudomonas aeruginosa* and *Neisseria gonorrhoeae* (Fig. 8B) (106, 107).

This structural homology, in addition to the protein sequence homology and biochemical characterization of ComGC oligomers, strongly suggests that ComGC oligomers are pili or pilus-like filaments. A wealth of pilus filament structural models have revealed that individual pilin N-terminal alpha helices make up the inner core of these filaments and provide much of the

structural stability, while the C-terminal domains comprise the solvent-facing outer surface of filaments (83). Therefore, ComGC C-terminal domain residues are the best targets for ComGC maleimide labeling, as they are mostly likely to be both solvent-exposed and non-essential for pilus stability.

Production of a comGC cysteine codon substitution panel

Fortuitously, the ComGC C-terminal domain is only 43 residues long, which made cysteine substitution of each residue feasible. We employed the stitching-by-overlap-extension (SOE) PCR technique to rapidly generate all 43 *comGC* cysteine codon substitution mutant alleles (hereafter referred to as *comGC*^{Cys}) (108). Briefly, this technique uses two overlapping/mutagenic primers at the desired codon and two regular primers flanking each of the other primers to create two PCR products. These PCR products both contain the desired mutation, as well as homology to one another around the mutated codon. Both products are added to a standard PCR, prime off one another via the homologous bases, and are extended to make a full-length gene that includes the desired mutation. The outermost primers are added after fifteen cycles, which then allows for exponential amplification of the desired mutant gene.

The *comGC*^{Cys} mutants were integrated into the *B. subtilis lacA* ectopic locus by natural transformation, and their expression placed under control of the native *comG* promoter and native *comGC* translation signals to ensure ComGC^{Cys} levels closely matched native ComGC^{WT} levels. At this stage, each strain contained one copy of both *comGC*^{WT} at the native *comG* locus, and *comGC*^{Cys} at *lacA*. We subsequently deleted native *comGC* through natural transformation of a Δ*comGC* PCR product, making *comGC*^{Cys} the sole copy of *comGC* in these strains (Fig. 9A) (109). PCR tests verified the insertion of *comGC*^{Cys} at *lacA* and the deletion of native *comGC* for each mutant. Sanger sequencing of the *comGC*^{Cys} PCR products confirmed the successful

creation of 41/43 of the intended mutants, with only *comGC*^{P64C} and *comGC*^{V75C} failing. We verified that our *comGC*^{Cys} ectopic expression construct could complement for Δ*comGC*^{WT} to restore transformability, as Δ*comGC* mutants are non-transformable (77). One of the 41 *comGC*^{Cys} mutants, *comGC*^{C81C}, expresses *comGC*^{WT} from the *lacA* ectopic site which should fully complement for native Δ*comGC*^{WT}. Surprisingly, this strain exhibited transformation efficiency orders of magnitude below wildtype, indicating a serious flaw in our complementation construct design (Fig. 9A).

The simplest explanation for this issue lies with the Δ*comGC*^{WT} allele. Rather than simply removing the native *comGC* ORF, this allele replaces the ORF with a kanamycin resistance cassette for convenient selection of mutants (Fig. 9A) (109). The *kan* resistance gene is expressed from a strong, constitutive promoter that also transcribes downstream genes within operons to counteract polar effects caused by cassette insertion (109). However, this can also drastically increase transcription of those downstream genes (109). In our case, this could cause a significant stoichiometric imbalance in *comGC* expression (expressed ectopically from the native *comG* promoter) compared to *comGD-G* (expressed potentially from both the native *comG* promoter and the *kan* promoter). ComGC-G are all pilin homologues, and altered stoichiometry of pilins can significantly affect pilus biogenesis in other systems (110–112). Therefore, *comG*-derived pilus biogenesis could be inhibited in our complementation system via overexpression of *comGD-G*.

To eliminate this possibility, we used the Cre-Lox recombination system implemented into the Δ*comGC*^{WT} allele (109). This system uses the site-specific recombinase Cre to remove the kanamycin resistance cassette in Δ*comGC*^{WT} via recombination of *lox* sites that flank the cassette, which consequently removes the *kan* promoter likely disrupting normal expression of

comGD-G (Fig. 9B) (109). We conjugated pDR244, a temperature-sensitive plasmid constitutively expressing *cre*, into the Δ*comGC*^{WT} ; *lacA::P*$_{comG}$*-comGC*^{WT} strain. After selecting for transconjugants and allowing for *lox* recombination, the plasmid was cured by growth at non-permissive temperatures, and single colonies were screened for loss of the antibiotic resistance markers associated with both the plasmid and kanamycin resistance cassette. Ectopically expressed *comGC*^{WT} was now able to fully complement for the recombined Δ*comGC*^{WT} allele, indicating that the original Δ*comGC*^{WT} allele inhibited transformation (Fig. 9B). The remaining 40 *comGC*^{Cys} complementation mutants were then put through the same Cre-Lox recombination, with recombination verified by PCR, to make the final panel of mutants used in this study.

ComGC^{Cys} *pilins support natural transformation*

We next decided to screen the *comGC*^{Cys} mutants for transformability, as non-transformable mutants would presumably be either unable to polymerize ComGC^{Cys} or produce defective polymers. Non-transformable mutants would, therefore, not be useful for studying biologically relevant ComGC^{Cys} structures. To this end, we cultured each mutant until naturally competent, transformed them with gDNA conferring resistance to spectinomycin, and then plated each transformed culture onto selective plates. A strain was considered transformable if at least one single colony resulted from transformation. In total, 16 mutants were transformable to some degree, with the majority of mutants having cysteine mutations in the first half of the C-terminal domain (Table 1 and Sup. Fig. 1). Interestingly, the most transformable mutants had cysteine mutations clustered close to the C-terminal endogenous cysteine (C81) in ComGC, which will be discussed below.

To better understand how well each transformable mutant transformed DNA, we performed quantitative transformation efficiency experiments. These experiments mirrored the initial

screens, but appropriate dilutions for each strain were made prior to plating on both selective and non-selective plates. We could then calculate how many transformants per CFU resulted from a transformation, giving us an objective measurement of transformation for each strain. Of these mutants, 5/16 had transformation efficiencies within 50-fold of wildtype, while 9/16 were within 500-fold (Fig. 10A). To give additional context as to the putative location of these mutated residues within ComGC, the residues have been colored green on the AlphaFold2 predicted ComGC structure presented in Fig. 8 (Fig. 10B). These 9 candidates were thus selected for further investigation in subsequent studies.

Discussion

Evidence mounts for ComGC-based pili in B. subtilis with monomer orientation within filaments homologous to well-characterized Gram-negative pili

The sequence homology of the ComGC N-terminal domain to a wide variety of structural pilins from both T2SS and T4P systems, the membrane localization of monomeric ComGC and the cell wall localization of oligomeric ComGC, as well as the conservation of a pre-pilin peptidase homologue *comC* along with a pre-pilin signal sequence within ComGC bearing the glutamate residue at position 5 essential for pilus production in other systems, all indicate that ComGC likely oligomerizes into pili or pilus-like filaments during *B. subtilis* natural competence (see Ch. 1 and Fig. 8) (76, 78, 80, 83, 105). The finding of *bona fide* pili in *S. pneumoniae* composed of homologous ComGC, however, is the best evidence in support of this hypothesis (93, 97). Our *in silico* prediction of monomeric ComGC's structure, generated using the Google Colaboratory-based AlphaFold2 software package, demonstrates that ComGC most likely folds into a traditional pilin structure with homology to known structural pilins from both *Pseudomonas aeruginosa* and *Neisseria gonorrhoeae* pilus systems (Fig. 8). This provides additional, strong support for the production of ComGC-based pili in *B. subtilis* during natural competence, when *comG* is expressed (11).

Critically, these data helped us target which residues within ComGC to mutate to cysteine for eventual fluorescent labeling via the maleimide labeling method (95). These residues should be both solvent-exposed and minimally perturbative to pilus formation to maximize the likelihood of fluorescent pilus labeling. Current experimental structural models of intact pili demonstrate (1) that pilin C-terminal domains comprise the outer filament surface in contact with the extracellular environment, and (2) that the bulk of inter-subunit molecular interactions stabilizing

pili are between the N-terminal alpha helices of adjacent pilin monomers found in the pilus core, including an essential salt bridge between the N-terminal amine of one pilin and the glutamate-5 sidechain (essential for pilus formation) of the neighboring pilin (83, 106, 107). We therefore decided the most logical placement of the cysteine substitution in ComGC would be within the C-terminal domain (colored blue in Fig. 8A), as determined by the AlphaFold2 structure.

Complementation of ΔcomGC via ectopic comGC^{Cys} expression suggests maintenance of pilin stoichiometry is essential for optimal ComGC pilus function

Ectopically expressed *comGC^{WT}* initially failed to complement for endogenous Δ*comGC* during natural transformation (Fig. 9A). We were intrigued by the possibility that Δ*comGC* was causing an imbalance in the normal stoichiometric ratios of ComGC to ComGDEFG, which may have interfered with ComGC pilus function (Fig. 9A). Ironically, the Δ*comGC* allele was cleverly designed with a strong constitutive promoter designed to express *kan* (for selection of transformants) and any downstream genes to prevent polar effects arising from *kan* cassette integration (109). In our case, however, this also had the potential to dramatically reduce the ratio of ComGC (produced using native transcription/translation signals) to ComGDEFG (produced using both native signals and the strong *kan* promoter). These fears were empirically founded, as the *kan* promoter in the deletion allele can increase transcription of downstream genes upwards of 200-fold within certain operons (109).

The pilin imbalance hypothesis was testable due to the *lox* sites flanking the *kan* cassette. We quickly removed the entire *kan* cassette (including the strong *kan* promoter), save for a 150 bp "scar", by introducing Cre recombinase into the complementation strains via conjugation of a constitutive *cre* expression plasmid (Fig. 9B). Ectopically expressed *comGC^{WT}* was able to restore full transformability in the resulting strain, demonstrating that the original *kan* cassette

impeded complementation via ectopic *comGC*^{WT} (Fig. 9B). The pilin imbalance hypothesis best explains this finding because the main alternative hypothesis, that the original Δ*comGC* allele caused polar effects on *comGD-G*, is extremely unlikely due to the presence of the *kan* promoter within the *kan* cassette.While the stoichiometric imbalance between the ComG pilins did appear to disrupt natural transformation, the mechanism of disruption remains unclear.

One possibility is that increasing the relative concentrations of ComGD-G to ComGC may create an artificially high number of pilus biogenesis initiation sites, effectively reducing the ComGC concentration at any given site, leading to short pili which cannot reach into the extracellular space or interact with environmental DNA. The role of minor pilins in initiating pilus biogenesis, as well as the modulation of pilus length based on minor pilin concentration, is well established (110–114). Additionally, while ComGD-G have not been observed within oligomers, each protein is essential for ComGC oligomerization, consistent with these pilins priming pilus biogenesis (77, 80). Therefore, this “ComGC dilution” hypothesis is plausible.Another equally plausible explanation is that one or more of the ComGD-G pilins have a role in mediating pilus retraction, so overproduction of these proteins may result in an increased likelihood of the retraction-mediator(s) being incorporated into elongating pilus filaments. This would consequently shorten the average length of pilus filaments, which could result in a similar situation noted above where pilus function is inhibited based on the inability of pili to extend properly. There are numerous examples of minor pilins being inhibitory to pilus extension and initiating pilus retraction, so this hypothesis is also plausible. Finally, it is worth noting that both hypotheses can be true simultaneously, with increased ComGD-G levels both diluting the available ComGC at a given initiation site and causing premature pilus retraction, which would only serve to compound the natural transformation deficiency.

Transformability of comGC^{Cys} complementation mutants paves the way for future fluorescent labeling of B. subtilis natural competence pili

The putative *B. subtilis* ComGC pilus presents unique challenges for implementing the maleimide labeling method (95). The presence of a pair of native cysteines within ComGC, necessary for both monomer stability and seemingly the production of oligomeric complexes, provides additional avenues for incompatibility with the cysteine substitutions required for maleimide labelling (80). BdbDC-mediated, off-target intramolecular disulfide bond formation between a native cysteine and the substituted cysteine would almost certainly impair the folding of ComGC^{Cys} monomers, potentially leading to degradation of ComGC^{Cys} and/or a pool of non-polymerizable monomers that could both dilute the pool of active ComGC^{Cys} and cause stalled extension of ComGC pili. Even for ComGC^{Cys} monomers that properly formed the intramolecular disulfide bond, the presence of the substituted cysteine could cause off-target intermolecular disulfide bond formation between ComGC^{Cys} monomers, potentially producing stalled pilus intermediates that are unable to continue extension.

Fortunately, these issues did not affect a subset of the *comGC^{Cys}* mutants. Approximately 40% (16/41) of the mutants were able to complement for native Δ*comGC* to some degree, with approximately 20% (9/41) resulting in transformation efficiencies within 500-fold of wildtype (Fig. 10 and Sup. Fig. 1). The top performing *comGC^{Cys}* mutant (*comGC^{A79C}*) was ~ 13% as efficient as wildtype, and the 2nd best performing mutant (*comGC^{V80C}*) ~ 3% as efficient (Fig. 10). This reduction is intrinsic to *comGC$^{A79C/V80C}$*, as full complementation was observed when *comGC^{WT}* was ectopically expressed from the same complementation background. Also of note is that both ComGC$^{A79/V80}$ are directly adjacent to ComGC^{C81} in the AlphaFold2 structure (Fig. 8A). This may suggest that many ComGC^{Cys} mutants suffer from the issues noted in the

preceding paragraph, but that ComGC$^{A79C/V80C}$, by virtue of having the substituted cysteine situated in close physical proximity to the endogenous ComGC cysteine, are able to tolerate off-target disulfide bond formation.

Looking to the future, our identification of these 9 transformable *comGC^{Cys}* mutants gives us the possibility of identifying *bona fide B. subtilis* ComGC pili for the first time via the maleimide labeling method and studying their production/disassembly dynamics and associations with DNA in real time. This will allow us to better understand the similarities and differences between homologous machineries responsible for DNA transport across the cell wall. We are planning on examining pilus production in strains expressing both *comGC^{WT}* and *comGC^{Cys}* alongside strains solely expressing *comGC^{Cys}*, potentially circumventing pilus production problems caused by ComGC^{Cys} by providing a more robust ComGC^{WT} scaffold within filaments that ComGC^{Cys} can incorporate into. Time will tell if maleimide labeling will work in *B. subtilis*.

Materials and methods

Bacterial strains and strain construction

All bacterial strains used in this study were derivatives of the prototrophic, domesticated *Bacillus subtilis* strain PY79 unless otherwise specified. General methods for strain construction were performed according to published protocols (115). Gene cloning to produce the *comGC*^{Cys} mutants in this study was performed by stitching-by-overlap-extension PCR (SOE PCR) using pKRH83 as template (108). Each *comGC*^{Cys} sequence was verified by Sanger sequencing. All mutants were produced via transformation of *B. subtilis* chromosomal DNA or PCR products into naturally competent PY79, except for the temperature sensitive *cre* expression plasmid pDR244 which was moved into PY79 strains via ICEBsI-mediated conjugation (116). Integration of both *comGC*^{Cys}, as well as the clean deletion of native *comGC*, were verified by PCR tests.

Media and growth conditions

B. subtilis strains were cultured in Lennox's Lysogeny Broth formulation (LB-Lennox) at 37 °C for general propagation and at 30 °C for temperature-sensitive plasmid conjugations, and Medium for Competence (MC medium: 1.07% (w/v) potassium phosphate dibasic, 0.52% (w/v) potassium phosphate monobasic, 0.088% (w/v) sodium citrate dihydrate, 0.0022% (w/v) ferric ammonium citrate, 2% (w/v) D-glucose, 0.1% (w/v) casein hydrolysate, 0.21% (w/v) L-glutamic acid monosodium monohydrate) at 37 °C for the consistent induction of natural competence and transformation. Cells were plated onto either nonselective Miller's Lysogeny Broth formulation (LB-Miller) 1.5% (w/v) agar plates or selective LB-Miller agar plates containing antibiotics MLS (macrolide-lincosamide-streptogramin: 1 μg mL^{-1} erythromycin + 25 μg mL^{-1} lincomycin), 10 μg mL^{-1} kanamycin, or 100 μg mL^{-1} spectinomycin as appropriate. During plasmid

conjugation, conjugations took place on Spizizen Minimal Salts plates (0.2% (w/v) ammonium sulfate, 1.4% (w/v) potassium phosphate dibasic, 0.6% (w/v) potassium phosphate monobasic, 0.1% (w/v) sodium citrate dihydrate, 0.02% (w/v) magnesium sulfate heptahydrate).

Natural transformation of B. subtilis

A single colony of each strain to be transformed was inoculated into 1 mL of MC medium + 3 mM $MgSO_4$ (pre-warmed to 37 °C) in 18 mm glass test tubes. The strains were cultured for 2.5 hr at 37 °C with 250 rpm orbital shaking so that each culture was in the mid-log growth phase. The cultures were concentrated by transferring to 1.5 mL microcentrifuge tubes, centrifuging for 5 min at 5,000 x RCF, removing all but ~ 200 μL supernatant, and then resuspending the cell pellets in the residual ~ 200 μL supernatant by repeated pipetting. Optical densities (OD_{600}) of 20-fold dilutions of the concentrated cultures (50 μL culture to 950 μL fresh MC) were determined using a spectrophotometer, and then the concentrated cultures were diluted to an OD_{600} = 0.05 into 1 mL fresh MC medium + 3 mM $MgSO_4$ (pre-warmed to 37 °C) in 18 mm glass test tubes to normalize cell density across all cultures. These new cultures were incubated for 3 hr at 37 °C with 250 rpm orbital shaking to induce natural competence. 200 μL of the naturally competent cultures were then transferred to 13 mm glass test tubes (pre-warmed to 37 °C), and saturating quantities of DNA (1 ng μL^{-1}) were typically added for transformation. The naturally competent cell/DNA mixtures would then be incubated for 2 hr at 37 °C with 250 rpm orbital shaking to allow for transformation to occur.

For strain construction, the transformations would be diluted 100-fold into PBS (2 μL transformation to 198 μL PBS), and then 200 μL of both undiluted and 100-fold diluted transformations would be plated onto selective LB-Miller agar plates and cultured overnight at 37 °C to obtain transformant single colonies. For quantitative transformation efficiency

experiments, chromosomal DNA from bJZ363 was used for transformation. The transformations would be serially diluted 10-fold down to 10^6-fold diluted. 200 μL of an appropriate dilution (giving easily countable colony numbers) would be plated onto selective LB-Miller agar plates, and 200 μL of 10^6-fold diluted transformations would be plated onto nonselective LB-Miller agar plates, and finally these plates would be cultured overnight at 37 °C to obtain transformant or total CFU single colonies, respectively. The total number of transformants and CFU would be calculated, and then transformation efficiency calculated by dividing transformants by CFU.

Modeling monomeric ComGC and manipulation of structures

The three-dimensional structure of monomeric *B. subtilis* ComGC was predicted using the Google ColabFold v1.5.3 webserver (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb) running AlphaFold2. The primary amino acid sequence of ComGC was obtained from SubtiWiki (http://subtiwiki.uni-goettingen.de), entered into the "query_sequence:" form of ColabFold, and all cells in the ColabFold notebook were run using default parameters (117). The highest ranked structure out of the five produced was selected as the putative structure. This structure was opened in Pymol, and all manipulations were done here, as well as all model renders found in Fig. 8 and Fig. 10. To render models, the ray trace command was used with resolutions of 1000 dpi, and the image dimensions set to 20 cm in height and enough width to capture the entire model at a given rotational state. These manipulations and renders were performed the same way for *P. aeruginosa* (PDB: 1OQW) and *N. gonorrhoeae* (PDB 2HIL) pilins.

Production of P_{comG}-comGC^{Cys} via stitching-by-overlap-extension PCR (SOE PCR)

Mutagenic primers for each desired comGC^{Cys} mutant (*comGC^{E56C}* – *comGC^{H98C}*) were designed with the aid of AAScan (118). The *comGC* coding sequence, along with 50 bases upstream and

downstream, was entered into the sequence field. The rest of the relevant parameters were entered as follows: “mut codon 1 and 2” = TGT and TGC; “min length” = 18; “max length” = 60; “MinGCclamp” = 2; “OptimizedGCclamp” box toggled on; “MinTm” = 56; “MaxTm” = 60; “MaxDeltaTm” = 2; “MinAnnealLen” = 15; “MinOverlap” = 15; “MaxOverlap” = 25.

One fragment within each *comGC*^{Cys} fragment pair to be stitched later was PCR amplified using the following protocol: 25 μL reactions consisting of (4 pg μL^{-1} pKRH83, 0.5 μM primer U-F + 0.5 μM primer C-F-XX (see supplementary strain construction protocols), 200 μM dNTPs, 3% (v/v) DMSO, 1X Q5 reaction buffer, 0.04 U μL^{-1} Q5 hot-start DNA polymerase (NEB)) with thermocycling conditions (1 cycle: 98 °C – 30 sec; 5 cycles: 98 °C – 10 sec, 62 °C – 30 sec, 72 °C – 2 min 40 sec; 25 cycles: 98 °C – 10 sec, 70 °C – 30 sec, 72 °C – 2 min 40 sec; 1 cycle: 72 °C – 2 min). The other fragment within each pair was amplified using the same protocol, swapping the primers for U-R and C-R-XX (see supplementary strain construction protocols).

Reactions containing fragments corresponding to the same *comGC*^{Cys} mutant were mixed, and residual pKRH83 was removed by directly adding 0.1 U μL^{-1} DpnI to the mixtures and incubating 37 °C – 2 hr. Residual primers were removed by magnetic bead cleanup using Sera-Mag SpeedBeads (Cytiva, 1:1 (v/v) bead mixture:PCR sample) according to the standard protocol for Sera-Mag Select beads. Purified fragment mixtures were then diluted 100-fold into 10 mM tris pH 8.0, and stitching reactions were performed as follows: 25 μL reactions consisting of (1:200 fragment mixtures, 200 μM dNTPs, 3% (v/v) DMSO, 1X Q5 reaction buffer) with thermocycling conditions (1 cycle: 98 °C – 30 sec; 15 cycles: 98 °C – 10 sec, 62 °C – 30 sec, 72 °C – 5 min 15 sec; 1 cycle: 72 °C – 2 min). Stitched products were amplified by adding a 1:1 (v:v) mixture of oJZ203:oJZ204 so both primers had concentrations of 0.5 μM, and then continuing thermocycling (1 cycle: 98 °C – 30 sec; 20 cycles: 98 °C – 10 sec, 72 °C – 30 sec, 72

°C – 5 min 15 sec; 1 cycle: 72 °C – 2 min). The outcome of these reactions was production of desired P_{comG}-$comGC^{Cys}$ mutants with homology flanks to the *lacA* locus, which allowed for integration into the *B. subtilis* genome via natural transformation and double crossover recombination.

Validating successful production of (lacA::P_{comG}-$comGC^{Cys}$; comGC::kan) strains

Chromosomal DNA was harvested from P_{comG}-$comGC^{Cys}$ *; comGC::kan* strains (bJZ140-170 ; bJZ172-183) as described above. To verify integration of *lacA::*P_{comG}-$comGC^{Cys}$, PCR of the mutated *lacA* locus was performed as follows: 25 μL reactions consisting of (4 ng μL^{-1} chromosomal DNA, 0.5 μM primer oJZ151 + 0.5 μM primer oJZ152, 200 μM dNTPs, 3% (v/v) DMSO, 1X Q5 reaction buffer, 0.04 U μL^{-1} Q5 hot-start DNA polymerase (NEB)) with thermocycling conditions (1 cycle: 98 °C – 30 sec; 30 cycles: 98 °C – 10 sec, 67 °C – 30 sec, 72 °C – 1 min; 1 cycle: 72 °C – 2 min). Samples were electrophoresed on 1% agarose in TAE gels, stained with 1X SYBR Safe stain (Invitrogen), and the products imaged on a BioRad ChemiDoc Touch instrument to verify that amplicons of the anticipated size were produced.

To verify deletion of *comGC* (*comGC::kan*), PCR of the mutated *comGC* locus was performed as follows: 25 μL reactions consisting of (4 ng μL^{-1} chromosomal DNA, 0.5 μM primer oJZ223 + 0.5 μM primer oJZ225, 200 μM dNTPs, 3% (v/v) DMSO, 1X Q5 reaction buffer, 0.04 U μL^{-1} Q5 hot-start DNA polymerase (NEB)) with thermocycling conditions (1 cycle: 98 °C – 30 sec; 30 cycles: 98 °C – 10 sec, 69 °C – 30 sec, 72 °C – 3 min 20 sec; 1 cycle: 72 °C – 2 min). Samples were electrophoresed on 1% agarose in TAE gels, stained with 1X SYBR Safe stain (Invitrogen), and the products imaged on a BioRad ChemiDoc Touch instrument to verify that amplicons of the anticipated size were produced.

Removing kan cassette from (P_{comG}-$comGC^{Cys}$; comGC::kan) strains via Cre-Lox recombination

Single colonies of original Δ*comGC* complementation strains (bJZ140-170 ; bJZ172-183) and the conjugative donor strain bearing the constitutive *cre* expression plasmid (bJZ447) were inoculated into 13 mm glass test tubes containing 0.5 mL of LB-Lennox or LB-Lennox + Spc^{50}, respectively, and then cultured for 16 hr at 30 °C with 250 rpm orbital shaking to generate saturated cultures (109). These cultures were diluted 50-fold (200 μL into 10 mL) into 125 mL Erlenmeyer flasks containing 10 mL of the same growth medium used in the first culturing step (pre-warmed to 30 °C), and then cultured for 2.5 hr at 30 °C with 250 rpm shaking to get cultures into the mid-log growth phase. Once complete, the OD_{600} of 1 mL each culture was determined using a spectrophotometer. These cultures were then diluted to OD_{600} = 0.02 into either 125 mL Erlenmeyer flasks containing 10 mL LB-Lennox (pre-warmed to 30 °C) for each *ΔcomGC* strain, or a 500 mL Erlenmeyer flask containing 50 mL LB-Lennox + Spc^{50} (pre-warmed to 30 °C) for the pDR244 strain to normalize cell density across each strain.

The strains were then cultured at 30 °C for 2 hr with 250 rpm shaking to get each culture to an OD_{600} = 0.2-0.3, at which point the conjugation machinery of the strain bearing pDR244 was induced by the addition of 1 mM IPTG and an additional 2 hr of culturing under the same conditions. The OD_{600} of each culture was then determined by measuring the OD_{600} of 10-fold diluted culture using a spectrophotometer so that the reading would be within the instrument's linear range. Each culture was concentrated to an OD_{600} = 20 by transferring to an appropriately sized centrifuge tube, centrifuging for 15 min at 25 °C with 4,300 x RCF, decanting the supernatant completely, and resuspending the cell pellets in an appropriate volume of 1X Spizizen Minimal Salts. Conjugative matings were then performed by mixing 150 of a *ΔcomGC* strain with 150 of the pDR244 strain, spotting 250 uL of each mixture onto 1X Spizizen Minimal Salt agar plates, letting the liquid dry/absorb into the plates, and then incubating the plates at 30

°C for 2 hr. The dense cell mixtures spotted on each plate were resuspended by repeatedly pipetting 5 mL of 1X Spizizen Minimal Salts onto the cells. Selected for transconjugants by plating 200 uL of each resuspended cell mixture onto LB-Miller + Spc^{100} + MLS plates and incubating at 30 °C for 18 hr.

The resulting transconjugant single colonies, having had grown for 18 hr, most likely produced the Cre recombinase and removed the kan cassette from Δ*comGC*. Next, pDR244 was cured by streaking a single colony of each transconjugant onto an LB-Miller agar plate, and then incubating the plates at 42 °C for 16 hr. Cells from the edges of 5 resulting single colonies per strain, having presumably resulted from the largest number cell divisions of any cells within the colonies and therefore being most likely to have lost pDR244, were patched onto the following plates: LB-Miller + MLS (screens for maintenance of P_{comG}-$comGC^{Cys}$ alleles, growth expected); LB-Miller + Kan^{10} (screens for recombination of Δ*comGC::kan*, growth not expected); LB-Miller + Spc^{100} (screens for loss of pDR244, growth not expected). Desired strains had the phenotype MLS^R; Kan^S; Spc^S, and strains conforming to this phenotype were selected for future experiments. Chromosomal DNA from each mutant was harvested, and recombination of Δ*comGC::kan* was verified by PCR using the same method described above for verifying integration of original Δ*comGC::kan* allele.

Chapter 3: Visualizing dynamic competence pili and DNA capture throughout the long axis of *Bacillus subtilis*

Adapted with permission from "Zuke JD, Erickson R, Hummels KR, Burton BM. 2023. Visualizing dynamic competence pili and DNA capture throughout the long axis of Bacillus subtilis . *J Bacteriol*"

Jason D. Zuke designed and carried out experiments, analyzed data, and wrote this chapter. Rachel Erickson analyzed data for this chapter. Katherine R. Hummels produced the *lacA::*P_{comG}*-comGC* construct that was vital for the production of this chapter. Briana M. Burton assisted with conceptualization and editing of this chapter and also gave general advice. Jonathan Lombardino and Tanya G. Falbel assisted with editing of this chapter.

Abstract

The first step in the process of bacterial natural transformation is DNA capture. Although long-hypothesized based on genetics and functional experiments, the pilus structure responsible for initial DNA-binding had not yet been visualized for *Bacillus subtilis*. Here, we visualize functional competence pili in *Bacillus subtilis* using fluorophore-conjugated maleimide labeling in conjunction with epifluorescence microscopy. In strains that produce pilin monomers within ten-fold of wild type levels, the median length of detectable pili is 300nm. These pili are retractile and associate with DNA. Analysis of pilus distribution at the cell surface reveals that they are predominantly located along the long axis of the cell. The distribution is consistent with localization of proteins associated with subsequent transformation steps, DNA-binding and DNA translocation in the cytosol. These data suggest a distributed model for *B. subtilis* transformation machinery, in which initial steps of DNA capture occur throughout the long axis of the cell and subsequent steps may also occur away from the cell poles.

Importance

This work provides novel visual evidence for DNA translocation across the cell wall during *Bacillus subtilis* natural competence, an essential step in the natural transformation process. Our data demonstrate the existence of natural competence associated, retractile pili that can bind exogenous DNA. Furthermore, we show that pilus biogenesis occurs throughout the cell long axis. These data strongly support DNA translocation occurring all along the lateral cell wall during natural competence, wherein pili are produced, bind to free DNA in the extracellular space, and finally retract to pull the bound DNA through the gap in the cell wall created during pilus biogenesis.

Introduction

Natural competence, the ability for bacteria to produce proteins that mediate the uptake of extracellular dsDNA, is widely distributed among the bacterial domain of life (10). DNA imported via natural competence is highly advantageous to cells and can be utilized in numerous, non-mutually exclusive ways. Once in the cytosol, the DNA can be metabolized to supply additional nucleotides and/or cellular carbon (119–121). If naturally competent cells encounter chromosomal DNA damage, the internalized DNA can be used for repair either through excision-mediated mechanisms or can be integrated into the resident chromosome via homologous recombination (122–125). Finally, the internalized DNA can confer novel genetic elements into the competent cell's genome (22, 126, 127). The process wherein DNA is internalized and subsequently integrated into the competent cell's chromosome is known as natural transformation and is one of the major drivers of horizontal gene transfer (HGT) in bacteria (128, 129).

The mechanisms governing Gram-positive natural transformation have been well characterized owing to work performed on model organisms such as *Streptococcus pneumoniae* and *Bacillus subtilis* (98). Transcription of genes involved in the production of DNA uptake and transformation proteins is activated through varied environmental cues (10). Once produced, a subset of these proteins must mediate the translocation of extracellular DNA across the negatively charged, 30-40 nm thick Gram-positive cell wall (36). The *comG* operon includes seven genes, each of which is homologous to components present in both type IV pilus (T4P) and type II secretion pseudopilus (T2SS) systems (130). These evolutionarily related systems employ a conserved set of proteins including at least one ATPase, at least one polytopic integral membrane protein, and a set of structural proteins known as pilins that work together to

dynamically assemble membrane-bound, filamentous protein helices composed of repeating pilin subunits (83, 131). These filaments are known as either pili for T4P, or pseudopili for T2SS, with the main discriminating characteristic being filament length. T4P pili can be multiple microns long and readily observed in the extracellular space, while T2SS pseudopili are typically not long enough to be visible in the extracellular space (86). Crucially, both systems can produce filaments which span the cell envelope and reach the extracellular space (T2SS pilin genes must be overexpressed for this to occur (86)).

The *comG* operon mediates the production of seven proteins that work together to form a T4P filament, composed of hundreds to thousands of individual ComGC pilin subunits, that extends through the cell wall into the extracellular space (80, 93, 97). Once extended, the ComGC pilus binds to free dsDNA in the environment (93, 97). Dynamic retraction of the ComGC subunits back into the membrane shortens the pilus, transporting bound DNA through the gap formed in the cell wall, into contact with the membrane-bound DNA receptor/binding protein ComEA (82, 132–134). After ComEA binding, DNA is brought into contact with the ComEC membrane channel (135, 136). The cytosolic protein ComFA aids in DNA entry across this channel via ATP hydrolysis (137–139). One strand of the incoming DNA is hydrolyzed, resulting in ssDNA in the cytosol (54, 57, 62). Through the action of various ssDNA binding proteins and the competence-specific RecA-loading protein DprA, RecA binds to the ssDNA and directs homologous recombination of the incoming ssDNA into the chromosome to complete natural transformation (140, 141).

Although this mechanism has been thoroughly researched, an interesting discrepancy remains regarding DNA translocation across the cell wall. While *S. pneumoniae* undoubtedly produces type IV pili to mediate initial DNA reception and translocation across the cell wall, all attempts

to identify such a structure in *B. subtilis* have yielded negative results, despite *B. subtilis* carrying a *comG* operon that is both essential for natural transformation and is homologous to the *S. pneumoniae comG* operon (77, 105). Biochemical data indicate that *B. subtilis* produces multimeric ComGC structures associated with the cell wall (80, 142). However, no microscopy data or structural studies have confirmed the existence of *bona fide* pili or pseudopili, so the exact nature of *B. subtilis* translocation of transforming DNA across the cell wall is unclear.

In this investigation, we report that ectopic expression of select *comGC* cysteine substitution alleles ($comGC^{Cys}$) in parallel with $comGC^{WT}$ allows for natural transformation to occur in *B. subtilis*. Expressing $comGC^{Cys}$ results in extracellular filaments that are labelable with a fluorescent maleimide-dye conjugate. We interpret these filaments to be pili composed mainly of ComGC. The pili can retract back towards the cell body and can bind to extracellular DNA. These data all support a model of DNA translocation across naturally competent *B. subtilis* cell walls by binding of DNA to pili, followed by pili retracting to transport bound DNA across the cell wall. We also find the localization of pilus biogenesis and DNA binding strongly biased to the periphery of the cell long axis. In addition, GFP-ComEA, and GFP-ComFA also reside predominantly along the long axis of the cell. Together the data suggest that DNA translocation across the cell wall is biased away from the poles, and even later steps in transformation may not be localized exclusively to or near the cell poles as has been previously reported (143, 144).

Results

ComGC cysteine substitution variants support transformation

Prior studies of natural competence associated pili, including the type IV competence pili of *V. cholerae* and *S. pneumoniae*, produced fluorescently labeled pili by using cysteine substitution variants of major structural pilins (96, 97). Since *B. subtilis* competence pilus biogenesis likely involves conversion of an intramolecular disulfide to an intermolecular disulfide between ComGC pilin subunits, we first assessed how the introduction of cysteine substitution variants affected transformability (80, 101). When *comGC*^{Cys} alleles were ectopically expressed as the sole copy of *comGC*, transformation efficiency dramatically decreased by multiple orders of magnitude compared to wild type (Sup. Fig. 2). When *comGC*^{WT} was expressed in the same strain background from the ectopic locus, however, a similar decrease in transformation efficiency was observed (Sup. Fig. 2E). This indicated to us that our ectopic expression construct had an intrinsic flaw that prevented proper complementation of the endogenous *comGC* deletion present in each strain. Importantly, this also implied that any strain ectopically expressing *comGC*^{Cys} that maintained transformation efficiency near that of the isogenic strain ectopically expressing *comGC*^{WT} was likely completely or mostly functional. Two alleles, *comGC*^{E56C} and *comGC*^{S65C}, allowed for transformation efficiency near that of *comGC*^{WT}, therefore we decided to focus on *comGC*^{E56C} and *comGC*^{S65C} going forward (Sup. Fig. 2CD).

To prevent any problems arising from deletion of endogenous *comGC*, we simply opted to ectopically express *comGC*^{E56C} or *comGC*^{S65C} without altering the native *comG* operon, resulting in strains co-expressing both *comGC*^{WT} and *comGC*^{E56C} or *comGC*^{S65C}. We surmised that any potential pilus would be able to incorporate both ComGC^{WT} and ComGC^{Cys}, allowing for pilus labeling via ComGC^{E56C} or ComGC^{S65C} subunits. These strains also carried an inducible *comK*

allele to increase the fraction of the population that would express late competence gene products and thus facilitate microscopic analysis (145). Both *comGC*^{E56C} and *comGC*^{S65C} had limited negative effects on transformation efficiency when co-expressed with native *comGC*. The strains expressing *comGC*^{E56C} and *comGC*^{S65C} under control of the native *comG* promoter were cultured until one hour from maximum competence, at which point *comK* expression was induced. Donor DNA carrying a spectinomycin resistance gene at a separate ectopic locus was transformed after 1 hour of induction. The strains producing the cysteine variants had transformation efficiencies within an order of magnitude of the matched wild type as well as a strain expressing only the endogenous copy of *comGC* (Fig. 11A). These transformation efficiencies were consistent with a functional apparatus for DNA translocation across the cell wall, which further encouraged us to continue using *comGC*^{E56C} and *comGC*^{S65C} in future experiments.

Next, we assessed ComGC pilin levels in the comGC^{WT}/comGC^{Cys} co-expression strains, as well as an isogenic control strain lacking an ectopic *comGC* copy. Previous investigations into type II secretion system pseudopili have demonstrated a direct relationship between pilin levels and pilus length (113). Therefore, we sought to verify that any pili observed would be biologically relevant, rather than artifacts caused by increased levels of pilin present in *comGC*^{Cys} mutants. Whole cell lysates of each strain were produced after inducing competence for one hour as described above and prepared for immunoblots with antiserum to detect ComGC (Fig. 11B). No ComGC was detected for an uninduced Δ*comK* culture, confirming ComGC-specific signal. ComGC levels for both strains carrying *comGC*^{Cys} at *lacA* were within ten-fold of the isogenic control strain carrying only the endogenous *comGC* copy (Sup. Fig. 3). We note that although there was much higher variability in ComGC levels in the co-expression strains expressing

comGC^{E56C} or *comGC*^{S65C} compared to strains with either a single or two copies of wild type *comGC,* these measurements support that any extended pilus identified is not simply due to excess ComGC subunits in the cell but is physiologically relevant.

Multiple ComGC cysteine substitution variants produce extracellular filaments

With multiple transformable *comGC*^{Cys} alleles identified, we next wanted to determine if strains expressing either allele produced pili observable in real-time via epifluorescence microscopy. To this end, we employed the widely used maleimide labeling method that has successfully identified natural competence-associated type IV pili from both Gram-negative and Gram-positive species (96, 97). This method utilizes the highly selective reactivity of maleimide towards thiol moieties. If a pilus is produced, and the major constituent pilin subunit contains an unpaired cysteine that faces the solvent, maleimide conjugates in the solvent will covalently bond with the free thiol group on each available ComGC^{Cys} cysteine residue. If the maleimide is conjugated to a fluorophore, epifluorescence microscopy can be used to identify fluorescent pili and observe their dynamic production and depolymerization in real-time.

To perform maleimide labeling, we first cultured *comGC*^{WT}, *comGC*^{E56C}, and *comGC*^{S65C} to competence as described above. Samples of each culture were incubated with Alexa Fluor 488 C_5 Maleimide in the dark, washed, and deposited on agar pads for microscopic analysis. Both *comGC*^{E56C} and *comGC*^{S65C} produced stubby, filamentous structures emanating from the cell periphery that resembled short pili (Fig. 12A). Critically, there was not a single instance of such structures being produced by *comGC*^{WT} treated with AF488-Mal (Fig. 12A and Sup. Fig. 4). This demonstrates that observation of these filaments depends on the presence of the ComGC^{Cys} monomers, lending credence to the idea that ComGC is incorporated into filaments that project into the extracellular space. Due to this finding, and the wealth of information indicating

ComGC's homology to other known pilins in terms of both form and function, we will henceforth refer to these filamentous structures as "ComGC pili".

The pilus production capacity of both strains was assessed quantitatively (Fig. 12B). Both intact pili that remained co-localized with the cell body, as well as regions of bright fluorescence near the periphery of the cell that we surmised were sheared pilus fragments (see discussion), were identified for 300 individual cells of each strain. While both strains produced both classes of signal, *comGC*^{S65C} consistently produced a greater number of pili and pilus-like features compared to *comGC*^{E56C}. We decided to move forward in our analysis using *comGC*^{S65C} based on these results. Qualitatively, these ComGC pili were notably shorter than competence pili analyzed for other species across the bacterial domain (96, 97). Therefore, we assessed the distribution of pilus lengths for comparison (Fig. 12C). The pili produced by both *comGC*^{E56C} and *comGC*^{S65C} strains were indeed quite short, with mean pilus lengths of 0.32 μm and 0.33 μm and standard deviations of 0.11 μm and 0.17 μm, respectively. The implications of this size distribution will be expanded upon in the discussion.

ComGC pili bind to DNA

One of the outstanding questions for *B. subtilis* DNA internalization is the mechanism by which transforming DNA is translocated across the cell wall (98). The prevailing hypothesis, that a pilus-like structure is formed which binds DNA and retracts to convey DNA to the membrane-localized DNA receptor ComEA, can now be tested directly with the aid of ComGC pilus maleimide labeling (98). First, we sought to address whether binding to DNA by ComGC pili occurs in the extracellular space during natural competence. To this end, an ~4.5 kb PCR fragment from the *ycgO* locus was produced and covalently labeled with Cy5 fluorophores along the length of DNA. This labeled PCR product can be co-incubated with maleimide labeled cells

to probe interactions between ComGC pili and extracellular DNA in real-time via time lapse epifluorescence microscopy (96, 97).

After silica column purification of the labeled PCR product, *comGC*^{S65C} was maleimide labeled as described previously. To ensure that we could visualize all possible pili around the periphery of the cell, we collected Z-stacks with a small step size (0.2 µm). Collecting the z-stacks also generated a time lapse of the cells with roughly five second intervals between images. In numerous instances labeled PCR product was already co-localized with pili at the onset of imaging, and in other cases migrated towards pili and subsequently remained co-localized with the pili (Fig. 13). Even subtle movements of pili were mirrored by co-localized PCR product over time. The persistent binding of PCR product through time, the coordinated movement of co-localized PCR product and pili, as well as the real-time migration and co-localization of PCR product to pili, would be highly unlikely to occur unless a *bona fide* binding interaction was being formed between pili and PCR product. These data provide direct evidence of the binding of DNA to *B. subtilis* ComGC pili.

ComGC pili are retractile

Once a ComGC pilus has formed and bound to extracellular DNA, a translocation step must ensue to move the bound DNA across the cell wall, where it can then presumably interact with the membrane-bound DNA receptor ComEA (82). Given that type IV pili are dynamic structures that can retract back into the membrane by the sequential disassembly of pilin monomers into the membrane, the most parsimonious mechanism of translocating DNA across the cell wall is simply depolymerizing the ComGC pilus (131). Maleimide labeling of ComGC pili allows us the opportunity to study the dynamics of pilus production and depolymerization, which allows us to verify whether DNA translocation is driven by ComGC pilus retraction.

To assess ComGC pilus retraction, *comGC*^{S65C} was maleimide labeled and deposited on agar pads as described previously. Time lapse imaging was performed using ten second intervals for one to two minutes to search for pili in the process of retraction (Fig. 14A). ComGC pilus retraction events were documented, wherein the length of a pilus would steadily decrease until only a strong focus of signal would persist at the cell periphery. For each pilus retraction event documented, the difference in pilus length between sequential frames was determined (44 total intervals across 10 pili). (96, 97)Interestingly, we observed two populations of retractile pili, one being considerably faster than the other. The median retraction rate for the "slow" group was 3 nm/sec, while the "fast" group was 12 nm/sec. These retraction rates were both notably slower than those calculated for competence pili of either *V. cholerae* or *S. pneumoniae*, which are much more in line with other observed type IV pilus systems (96, 97). All members within the fast population did not have statistically significant differences in average retraction rate nor did all members within the slow population (ANOVA, $p = 0.36$ and $p = 0.69$ respectively). However, a strong statistically significant difference was observed when the entire fast population and the entire slow population set of retraction rates were compared (T-Test, $p = 3.4 \times 10^{-7}$). These data will be examined further in the discussion section.

ComGC pili, DNA, ComEA, and ComFA localize along the long axis of competent cells

Our ability to identify sites of pilus biogenesis presents the opportunity to reassess the spatiotemporal dynamics of DNA uptake during natural competence. DNA uptake has been previously shown to occur primarily at or near the cell poles, perhaps taking advantage of a natural structural weakness in the cell wall region at the interface of the lateral and polar cell wall sections (143, 144). To probe this hypothesis, we analyzed the localization of ComGC pilus biogenesis in *comGC*^{S65C}, as well as the localization of stably bound fluorescent PCR product. If

DNA uptake occurs predominantly at or near the cell poles, we should expect the majority of assembled pili, and the majority of bound DNA, to localize close to the cell poles during natural competence.

The analysis of ComGC pilus localization, employing approximately 230 individual pilus biogenesis events, demonstrated that pili emanate along the long axis of competent cells at the cell periphery, with relatively few events occurring at or near the cell poles (Fig. 15A). In agreement with these data, fluorescent DNA that was stably bound to pili or the cell periphery primarily localized along the long axis of competent cells (Fig. 15B). To further assess this model of DNA uptake, we compared the distribution of ComGC pili to subcellular localization of functional GFP-ComEA and GFP-ComFA. Since ComEA serves the vital role of binding incoming DNA at the cell membrane, the protein should localize at the sites of initial DNA entry, which presumably coincides with the location of pili biogenesis. In accordance with this hypothesis, GFP-ComEA localizes along the long axis of competent cells, just like ComGC pili and bound DNA (Fig. 15C). There are two possibilities for the next phase of DNA entry. The handoff from ComEA to the ComEC channel could occur near the site of pilus biogenesis, or it could occur at a secondary location. ComFA, which powers DNA translocation across the cell membrane, thus should serve as a proxy for the cytoplasmic internalization step. Indeed, GFP-ComFA, like all the other competence proteins and DNA assessed above, localizes near the cell periphery along the long axis of competent cells (Fig. 15D). It would appear, therefore, that DNA translocation across the cell wall, and entry into the cytoplasm, need not occur proximal to a cell pole.

Discussion

In *Bacillus subtilis*, the *comG* operon is essential for natural transformation (77). The ComG proteins of *B. subtilis* have been assumed to be involved in the production of a pilus capable of conveying extracellular DNA across the cell wall. This idea stems from multiple lines of evidence: each protein coded for in the operon has homology to components of either type IV pilus or type II secretion pseudopilus systems, there is biochemical evidence supporting multimerization of ComGC associated with the cell wall, and more recently ComGC pili have been demonstrated in *Streptococcus pneumoniae* (80, 97, 130, 142). However, direct evidence of ComG-mediated pilus production in *B. subtilis* has been lacking, with no microscopy or structural studies supporting pilus biogenesis. Here we demonstrate that *B. subtilis* in the naturally competent state produces ComGC-based pili. These pili are capable of stable binding of DNA, and they can retract dynamically back towards the cell membrane after assembly. ComGC pili were also found to localize along the long axis of the cell at the cell periphery, mirroring the localization of bound DNA and the downstream essential competence proteins ComEA and ComFA we observed.

We demonstrated that $comGC^{E56C}$ and $comGC^{S65C}$ alleles were functional by complementation of *ΔcomGC* for transformation efficiency (Sup. Fig. 2). These $comGC^{Cys}$ variants complemented almost as well as $comGC^{WT}$, further supporting their functionality. The *comGC* complementation construct itself caused massive decreases in transformability regardless of the *comGC* allele used, and no pili were ever observed in either of the *ΔcomGC* complementation strains (Sup. Fig. 5). The most likely explanation for the poor transformability of these strains is a polar effect due the introduction of a sub-promoter into the *comG* operon during native *comGC* deletion (109). There are four additional pilin gene homologues downstream of *comGC* in the operon (75).

Overexpression of certain pilins can be inhibitory for pilus production, so it's feasible that the sub-promoter could cause overexpression of *comGDEFG*, and at least one gene product could inhibit pilus biogenesis (112). Regardless of the exact reason, the failure of the complementation strains to produce pili spurred us to try another approach to produce biologically relevant, visualizable pili.

Rather than alter the endogenous *comG* promoter, we instead took a less perturbative approach and expressed $comGC^{Cys}$ alongside native *comGC*. In stark contrast to the *comGC* complementation strains, we demonstrated that $comGC^{E56C}$ and $comGC^{S65C}$ co-expression strains closely approximated wild type transformation efficiencies (within two-fold, Fig. 11A). The levels of total ComGC in both $comGC^{E56C}$ and $comGC^{S65C}$ were variable and slightly lower than for the isogenic control (Fig. 11B). While this may suggest that the ComGC pili we observed were artifactually shortened due to a smaller ComGC pool, it must be noted that the ComGC peptide used to raise the ComGC antiserum we employed for Western blotting included both positions E56 and S65 (105). It is therefore possible that the ComGC antiserum may not recognize $ComGC^{E56C}$ or $ComGC^{S65C}$ as efficiently as $ComGC^{WT}$ since the binding epitope could have been altered. Total ComGC levels in $comGC^{E56C}$ and $comGC^{S65C}$ may therefore be roughly equivalent to the isogenic control, making the pili we observe reflective of those found in wild type cells, or possibly slightly shorter.

The mean pilus length measured in *B. subtilis* of 0.33 μm with simultaneous expression of $comGC^{WT}$ and $comGC^{S65C}$ (Fig. 12C) is notably shorter than the mean lengths of competence pili identified in *V. cholerae* or *S. pneumoniae* of 1 μm and 0.5 μm respectively (96, 97). The somewhat smaller difference between *B. subtilis* and *S. pneumoniae* ComGC pili may be attributable to pilus measurement error. Exact pilus start and end points had to be manually

assigned based on the phase-contrast and epifluorescence images taken, which could lead to length differences based on user interpretation. Additionally, we determined pilus start points to be at the border of the phase-contrast (cell body) signal, not the border of the epifluorescence (ComGC) signal. There was generally a gap between the internal ComGC signal and the external cell body signal, so pilus lengths would have been increased had pilus start points been marked at the ComGC signal border. Another likely source of variation could be the presence of endogenous cysteine residues in *B. subtilis* ComGC. An off-target, BdbDC-mediated disulfide bond between an endogenous cysteine on one ComGC monomer and the mutant cysteine of another ComGC^{Cys} monomer might result in early termination of pilus elongation due to conformational changes at the pilus base that inhibit proper coordination of the pilus biogenesis apparatus. *S. pneumoniae* ComGC, in contrast, contains no endogenous cysteine residues that could improperly form disulfide bonds with the ComGC^{Cys} cysteine (see PDB 5NCA), negating this possibility. And finally, the observed length difference may simply stem from intrinsic differences in the protein sequences of each ComG protein comprising the two systems.

The larger difference between *V. cholerae* and *B. subtilis* pilus lengths may stem from differences in how force for pilus biogenesis is generated. *V. cholerae* employs both a dedicated extension (PilB) and a retraction (PilT) ATPase, whereas *B. subtilis* only has one identifiable pilus ATPase homologue (ComGA) (96, 130). PilB ATPase activity may simply be faster and/or more processive than that of ComGA, which would grant PilB greater ability to incorporate new pilin subunits into a growing pilus prior to disassembly than ComGA, making the average pilus longer for PilB-polymerized pili. Further investigation of pilin levels and extension/retraction ATPase activities are warranted to produce a complete picture of how competence pilus lengths are established.

Our work provides evidence of direct *B. subtilis* ComGC pilus-DNA interactions with physiological levels of ComGC proteins (Fig. 13). Such interactions are consistent with data from a diverse set of naturally competent bacteria, including *V. cholerae*, *S. pneumoniae*, *Neisseria gonorrhoeae*, and *Thermus thermophilus* that demonstrate either direct binding of pili to DNA or individual pilins to DNA (92, 96, 97, 146). We also demonstrate that *B. subtilis* ComGC pili are retractile in nature, as has been observed for *V. cholerae* and *S. pneumoniae* competence pili (Fig. 14) (96, 97). Intriguingly, we observed two distinct populations of retracting pili with variable retraction rates: a slow population that retracts with a median rate of 3 nm/sec and a fast population that retracts with a median rate of 12 nm/sec (Fig. 14C). The slow population may be retracting spontaneously, whereas the fast population is most likely actively retracting via ComGA activity. Spontaneous retraction of *V. cholerae* competence pili has been observed with a notably slower retraction rate compared to active retraction, strengthening this idea (96, 147). Moreover, the clear distinction in retraction rates between the two populations suggests that a specific process increases the retraction rate of the fast-retracting pili. The most parsimonious explanation is, of course, that the pilus ATPase homologue ComGA is promoting active retraction events within the fast population.

The retraction rate for the fastest ComGC pili observed (median = 12 nm/sec) is much slower than for *V. cholerae* (median ~ 100 nm/sec) and *S. pneumoniae* (median ~ 80 nm/sec). We consider two possible explanations for this difference. First, the rate differences could be reflective of differences in the enzymatic activities of the ATPases utilized for retraction across these systems (Fig. 14B) (131). Alternatively, the potential disulfide isomerization involved in assembly and disassembly of ComGC pili in *B. subtilis* may impact the retraction rate. Dissection of the contribution of the disulfide is not trivial, as disulfide bond formation is

necessary for ComGC pilus biogenesis (80), but future mechanistic studies to address this question will be valuable.

Subcellular localization analysis of ComGC pili and of associated DNA produced a localization pattern quite distinct from the previously reported polar localization patterns of other late competence gene products, including ComGA and ComEC (143, 144). We found that ComGC pili and associated DNA both predominantly localize across the long axis of the cell (Fig. 15AB), with little clustering of ComGC pilus biogenesis. Additionally, ComEA and ComFA were seen to localize along the long axis of competent cells (Fig. 15CD). These localization patterns are consistent with the predominant distributions reported for DNA bound at the cell membrane (133). These observations suggest that the initial step of mobilizing extracellular DNA across the peptidoglycan layer for ComEA binding at the cell membrane may occur throughout the surface of the cell, predominantly along the long axis, and may even be excluded from the cell poles. One possible explanation for this is that during natural competence the lateral cell wall is mostly static, as the cells have ceased growth and elongation, and any remaining peptidoglycan remodeling is likely occurring at the cell poles at sites of cell separation (148–150). This may facilitate ComGC pilus formation at this region, since steric clashes with active cell wall remodeling systems could potentially be reduced, and any channels formed in the cell wall may be more likely to remain open for extended periods of time.

While our data for ComEA localization are largely consistent with previous reports, our analysis differs significantly for ComFA localization (133, 143). We observed GFP-ComFA puncta throughout the long axis of the cell (Fig. 15D), whereas polar ComFA-YFP localization had been observed previously. This difference is of critical importance since the localization of ComFA to the cell poles, as well as that of other late competence proteins such as ComEC and ComGA, led

to the conclusion that DNA translocation across the cell membrane occurred at or near the cell poles (143). Our data, on the other hand, are consistent with DNA entry points distributed along the lateral cell membrane.

To address the disparities, we consider DNA uptake in the context of the two-stage model for transformation (151). First, our data may best represent the location of the first step of DNA uptake, where DNA crosses the cell well. The localization of pilus biogenesis and retractile pili that interact with DNA provides a direct view of active competence pili. In contrast, the polar localization of ComGA holds the implicit assumption that the foci are coincident with active protein. (143, 152). Given the necessity of ComGA for pilus assembly, functional ComGA should localize to the sites of pilus biogenesis (i.e. along the cell long axis), but that has not been observed (142). This raises questions as to the biological relevance of the observed localization patterns of the late competence protein fluorescent fusions used in previous studies. Subpopulations of proteins are often sufficient for biological activity, and active fractions do not always correspond to the visualized population (153). This could be the case for the prior results with ComGA fusions.

The location of the second stage of DNA import, across the cell membrane as assessed by ComEC and ComFA localization, remains less clear. Studies from multiple groups have come to the same conclusion regarding polar localization of ComFA (143, 152). The constructs used in the respective localization experiments differ from ours in placement of the fluorescent tag. The data presented here use a complementing amino-terminal fluorescent fusion to ComFA. Consistent with prior published work, we observe predominantly polar localization when the fluorescent protein is at the carboxy-terminal end of ComFA. However, there is significant proteolysis of the fluorescent tag and free ComFA in these cells, reducing confidence in these

fusions as reporters of active protein. Thus, we favor the amino-terminal fusion presented here as reporting on functional localization. However, from these data alone we cannot rule out that DNA translocation across the cell membrane could occur at cell poles. In such a system, DNA that had been captured by ComEA after translocation across the cell wall would most likely be transported to the cell pole, where the DNA would then be internalized (154).

Critically, our results expand on the mechanism of DNA reception and translocation across the cell wall during Gram-positive natural competence, where models have thus far relied on observations made solely using *S. pneumoniae* as a model organism. Our observations are most consistent with the following model of DNA translocation across the cell wall. During natural competence, the protein products of the *comG* operon work together to generate an extracellular pilus that is comprised primarily of ComGC along the long axis of the cell. These pili are dynamic and retract stochastically. At some point, dsDNA in the environment binds to the pilus surface. At some point after DNA binding, the pilus will retract back into the membrane, which will consequently pull the bound DNA through the gap left in the cell wall. Once across the cell wall, the DNA will bind to ComEA at the cell membrane, and translocation into the cytoplasm will commence, either at a cell pole or along the long axis of the cell.

Numerous questions remain unresolved regarding the mechanism of DNA translocation across the cell wall. Are each of the 5 pilin homologues in the *comG* operon present in the competence pilus, and if so, where are they located? Each gene in the *comG* operon is essential for transformation, so presumably each pilin is involved to some degree with pilus biogenesis and/or function (77). In *S. pneumoniae*, it was recently discovered that ComGC, ComGF, and ComGG were present throughout pili, although no other pilins were detected (155). It is possible that the other pilins (ComGD and ComGE) are located at the tip of the pilus in single copy and are

responsible for DNA binding. This is consistent with the role of low abundance “minor pilins” in initiation of pilus assembly and interactions with the environment (114, 131, 155). Identification of minor pilin point mutations that allow for DNA-binding-deficient pili to be produced, as was achieved for *V. cholera* competence pili, would support this hypothesis (96). With successful methods to visualize ComGC pili and DNA capture in *B. subtilis*, several of these open questions now become accessible.

Materials and methods

Strain construction

General methods for strain construction were performed according to published protocols (156, 157). Molecular cloning was performed using the Gibson assembly method with HiFi assembly enzyme mix (NEB) (158). PCR amplification templates were derived from either Chromosomal DNA isolated from the prototrophic domesticated *B. subtilis* strain PY79 or plasmid pKRH83. Introduction of DNA into PY79 derivatives was conducted by transformation (159). The bacterial strains, plasmids, and oligonucleotide primers used in this study are listed in Supplementary Tables 4-6 respectively.

Media and growth conditions

For general propagation, *B. subtilis* strains were grown at 37°C in Lennox lysogeny broth (LB) medium (10 g tryptone per liter, 5 g yeast extract per liter, 5 g NaCl per liter) or on LB plates containing 1.5% Bacto agar. Where indicated, *B. subtilis* strains were grown in the nutrient-limiting medium Medium for Competence with 2% fructose (MC-Fru ; 61.5 mM K_2HPO_4, 38.2 mM KH_2PO_4, 2% (w/v ; 110 mM) D-fructose, 3 mM $Na_3C_6H_5O_7 \cdot 2H_20$, 80 µM ferric ammonium citrate, 0.1% (w/v) casein hydrolysate, 11 mM L-Glutamic acid potassium salt monohydrate) substituted for 2% glucose to prevent catabolite repression of P_{xyl} promoter (156). When appropriate, antibiotics were included in the growth medium as follows: 100 µg mL^{-1} spectinomycin, 5 µg mL^{-1} chloramphenicol, 5 µg mL^{-1} kanamycin, 10 µg mL^{-1} tetracycline, and 1 µg mL^{-1} erythromycin plus 25 µg mL^{-1} lincomycin (mls). When required, 0.5% (w/v ; ~30 mM) D-xylose was added to cultures to induce protein expression.

Producing lysed protoplasts for B. subtilis transformation

A single colony of the *B. subtilis* strain bBB050 (Cm^R) was inoculated into LB medium and incubated at 37 °C with 250 rpm shaking for 3 hr. The OD_{600} of a 1:10 dilution of the culture was measured, and then 1 mL of culture was pelleted at 21,000 x G – 2 min and the supernatant removed completely. The pellet was resuspended to an OD_{600} = 10 in *Bacillus* protoplasting buffer (50 mM tris pH 8.0, 50 mM NaCl, 5 mM $MgCl_2$, 25% (w/v) sucrose, 0.2 mg/mL lysozyme) and incubated in a 37 °C water bath for 30 min to protoplast cells. The sample was removed from the water bath and left at room temperature until transformation, at which time protoplasts were pelleted at 10,000 x G for 5 min, the supernatant was completely removed, and then an equal volume of ddH_2O was added. The protoplasts were lysed by resuspension in the ddH_2O by repeated pipetting.

Transformation efficiency assays

Single colonies of *B. subtilis* strains of interest were inoculated into MC-Fru and cultured at 37 °C with 250 rpm shaking until an OD_{600} = 0.2 – 0.5 was reached. Cells were pelleted at 10,000 x G – 2 min and resuspended in ~15% of residual supernatant to concentrate cells, OD_{600} of a 1:20 dilution of resuspended cells was measured, and the resuspensions were diluted into fresh MC-Fru to an OD_{600} = 0.05. Cultures were incubated at 37 °C with 250 rpm shaking for 2 hr (1 hr prior to max natural competence induction in these conditions), and strains containing P_{xyl}-*comK* were induced with 0.5% xylose to maximize the proportion of competent cells in the populations. The strains continued to be cultured at 37 °C with 250 rpm shaking for 1 hr to allow for maximal natural competence induction, and then ~ 10^5 lysed Cm^R protoplasts per µL competent cells were added to the cultures (139). Transformation was allowed to proceed under the same culturing conditions for 2 hr. 10-fold serial dilutions of each sample were made down to 10^6-fold diluted in PBS, and appropriate dilutions were plated onto LB and LB – Cm^5 agar plates and incubated for

16-20 hr at 37 °C to allow for colony growth. The number of transformants and total cells were calculated from the single colonies on LB – Cm^5 and LB respectively, and the transformation efficiency was calculated as the ratio of transformants to total cells in a given sample.

ComGC Western blotting

B. subtilis strains of interest were cultured according to the same protocol noted for transformation efficiency assays, but 1 mL of culture was centrifuged at 21,000 x G for 2 min to pellet cells after 3 hr incubation post-dilution, and supernatant was completely removed. The OD_{600} of a 1:10 dilution of each culture was measured, and each cell pellet was resuspended in cell lysis buffer (25 mM Tris pH 8.0, 25 mM NaCl, 3 mM $MgCl_2$, 1 mM $CaCl_2$, 0.2 mg/mL lysozyme, 0.1 mg/mL DNase I) to an OD_{600} = 10 based on the previous measurements. Resuspended cells were incubated in a 37 °C water bath for 20 minutes to lyse cells and degrade genomic DNA. Cell lysates were mixed with an equal volume of 2X reducing tricine sample buffer (200 mM tris pH 6.8, 40% (v/v) glycerol, 2% (w/v) SDS, 0.04% (w/v) Coomassie Blue G-250, 2% (v/v) beta-mercaptoethanol) and heated to 37 °C for 30 minutes for protein denaturation. 5 µL of each preparation was added to wells of a tris-tricine mini gel (Stacking gel: 1M tris pH 8.45, 4% (w/v) acrylamide/bis-acrylamide (29:1) ; Resolving gel: 1M tris pH 8.45, 15% (w/v) glycerol, 10% (w/v) acrylamide/bis-acrylamide (29:1)) and electrophoresed (Running buffer: 0.1 M tris-Cl, 0.1 M tricine, 0.1% (w/v) SDS) at 100 V until loading dye exited the gel (typically 1.75 – 2 hr). Separated proteins were Western transferred to a 0.2 µm PVDF membrane using the semi-dry transfer method at 15V for 20 min with Towbin transfer buffer (160). The membrane was washed 3X in ddH_2O for 5 min with gentle shaking to remove transfer buffer, then the membrane was stained with 0.1% Ponceau S solution to detect total protein for loading normalization. The membrane was blocked with 5% (w/v) milk in TBS-T for 1 hr with gentle

shaking, and then the membrane was incubated with 1:2000 rabbit-derived antisera containing 1° ComGC antibody (TBS-T + 1% (w/v) milk) overnight at 4 °C (77). The membrane was washed 3X in TBS-T for 5 min with gentle shaking, and then the membrane was incubated with 1:20,000 Abcam goat anti-rabbit 2° HRP antibody (TBS-T + 1% (w/v) milk) for 1 hr at room temperature with gentle shaking. The membrane was washed as described previously, and then developed using Clarity ECL substrate (Bio-Rad) and imaged for chemiluminescence using a Bio-Rad ChemiDoc Touch imaging system.

Preparing cover glass and agar pads for use in microscopy

All cover glass used in microscopy experiments was pre-cleaned prior to use. 22 mm x 22 mm #1.5 borosilicate coverslips (DOT Scientific) were placed in a Wash-N-Dry™ coverslip rack (Sigma) and submerged in ~80 mL of 1 M NaOH in a 100 mL glass beaker. The beaker was then placed into an ultrasonic cleaning bath (frequency = 40 kHz, power = 120W) and sonicated for 30 min to remove the thin grease layer present on the coverslips. The coverslip rack was submerged into a fresh 250 mL glass beaker filled with ddH_2O, the ddH_2O was removed, and then the coverslips were washed 3X with 250 mL of ddH_2O. The coverslips were then either air-dried overnight or dried immediately with compressed air.

For the preparation of agar pads, pre-cleaned borosilicate glass microscope slides were first rinsed free of detritus using ddH_2O and were then either air-dried overnight or dried immediately with compressed air. Working in a fume hood, half of the total rinsed slides were dipped into a 2% (v/v) solution of dichlorodimethylsilane (in chloroform) so that their entire surface was contacted by the solution. The solution was allowed to drip off the slides into the original container, and the remaining organic solvent on the slides was allowed to evaporate in the fume hood for 30 min. The slides were then thoroughly rinsed in ddH_2O and dried as mentioned

previously to generate dry glass slides with extremely hydrophobic surfaces. Two pieces of lab tape (VWR #89098-074) were placed on top of one another, running lengthwise across an unrinsed microscope slide. For every agar pad to be produced, 2 of these taped slides were made. ~15 minutes prior to imaging, conditioned MC-Fru medium (0.22 µm PES filter sterilized) from cultures grown to competence via the transformation efficiency assay protocol was heated to 37 °C, added 1:1 to 90 °C 2% (w/v) molten LE agarose (SeaKem) in ddH_2O, and then vortexed to make molten 1% (w/v) agarose in 0.5X conditioned MC-Fru medium. 20 µL of this mixture was applied to the surface of a hydrophobic glass slide, and a solidified pad was produced according to a previously established protocol, using a rinsed (but untreated and hydrophilic) glass slide to form the top of the pad (161). This specific setup allows for the agar pad to stick to the bottom slide and easily slide out from the top slide, leaving an unmarred and flat surface for imaging.

Fluorescent labeling – ComGC pili

For all experiments involving ComGC pilus labeling, *B. subtilis* strains of interest were cultured according to the same protocol noted for transformation efficiency assays. After 3 hr incubation post-dilution, 100 uL of each culture was transferred to a 37 °C pre-warmed 13 mm glass test tube, and 25 ug/mL Alexa Fluor 488-maleimide was added to each culture aliquot. The aliquots were incuabted in the dark at 37 °C on a rolling drum for 20 minutes to allow for ComGC^{Cys} pilin labeling. The aliquots were transferred to centrifuge tubes, centrifuged at 5,000 x G for 30 seconds to gently pellet cells, and all supernatant was removed. The cell pellets were washed by gently resuspending in conditioned MC-Fru medium (0.22 µm PES filter sterilized) via gentle pipetting. The cells were centrifuged again at 5,000 x G for 30 seconds, the supernatant removed, and the cells gently resuspended in one-tenth the original volume of conditioned MC-Fru medium.

Fluorescent labeling – PCR product

An ~4.5 kb PCR product was amplified from genomic DNA of a *B. subtilis* strain bearing P_{hysp}-*bdbDC* at the *ycgO* locus (bJZ185) using LongAmp Taq DNA Polymerase (NEB) and oligonucleotide primer pair (oJZ436 + oJZ437). This PCR product was purified using the E.Z.N.A. Cycle Pure Kit (Omega Bio-Tek) following the manufacturer's instructions. 1 µg of PCR product was fluorescently labeled using the *Label* IT-Cy5 Nucleic Acid Labeling Kit (Mirus) following the manufacturer's instructions, with a sufficient quantity of *Label* IT-Cy5 reagent to covalently link a fluorophore to between 1 in 20 and 1 in 60 bases (75 – 230 Cy5 molecules per DNA molecule). The Cy5 labeled PCR product was purified using the E.Z.N.A. Cycle Pure Kit as mentioned previously. The final product was electrophoresed on a 1% agarose in TAE mini gel at 100 V for 45 min, stained using SYBR Safe dye (Invitrogen) according to the manufacturer's instructions, and quantified by densitometric analysis of the product band compared to a reference band of known quantity.

Microscopy – Alexa Fluor 488-maleimide labeled ComGC pili and Label IT-Cy5 labeled PCR product

For the initial identification of ComGC pili, quantification of pilus production, pilus length measurements, and the determination of ComGC pilus retraction rates, cells were labeled with Alexa Fluor 488-maleimide as previously described. 0.5 µL of resuspended, labeled cells was applied to the center of a conditioned MC-Fru agar pad, and a pre-cleaned 22 mm x 22 mm #1.5 borosilicate coverslip was applied to the drop of culture. The coverslip was gently compressed with a gloved finger to ensure cells made contact with the agar pad, and then the space between the coverslip and glass slide was sealed by applying molten (60 °C) VaLAP sealant (1:1:1 petroleum jelly:lanolin:paraffin) to the coverslip edges. The cells were imaged using a Zeiss

Axio Observer.Z1 inverted epifluorescence microscope equipped with a Zeiss Plan-Apochromat 100X/1.4 Oil PH3 objective lens and and a Teledyne Photometrics CoolSNAP HQ^2 CCD camera. Cells were typically exposed to light from a Zeiss Colibri 469 nm LED module with Zeiss filter set 38 HE for 2 s at 100% LED power to observe ComGC pili. Cell bodies were imaged using phase-contrast microscopy (50 ms exposures). Imaging was performed at 37 °C, achieved by a PECON Heater S objective heater calibrated via thermocouple. Time lapse microscopy was performed with the above conditions, with exposures occuring every 5 – 11 s depending on the experiment.

For co-localization experiments on ComGC pili and PCR products, cells were labeled with Alexa Fluor 488-maleimide as previously described. Just prior to deposition of labeled cells onto a conditioned MC-Fru agar pad, Cy5 labeled PCR product was added to the cell suspension to a final concentration of 120 pg/uL. The new mixture of Alexa Fluor 488 labeled cells and Cy5 labeled PCR product was applied to a conditioned MC-Fru agar pad as described above. imaging was performed using a Zeiss Axio Imager.Z2 upright epifluorescence microscope equipped with a Zeiss Plan-Apochromat 100X/1.4 Oil PH3 objective lens and a Teledyne Photometrics Prime 95B sCMOS camera. Cells were typically exposed to light from a Zeiss Colibri 469 nm LED module with Zeiss filter set 38 HE for 200 ms at 20% LED power to observe ComGC pili, while *Label* IT-Cy5 labeled PCR product was observed using light from a Zeiss Colibri 631 nm LED module with Zeiss filter set 90 HE for 500 ms at 20% LED power. Cell bodies were imaged using phase-contrast microscopy (100 ms exposures). Z-stacks were acquired in 0.2 µm increments starting 1 µm above the midcell plane and progressing until 1 µm below the midcell plane. Due to a fortuitous imaging delay during phase-contrast acquisition, the time interval between epifluorescent image acquisition was ~ 5 s, producing a time lapse.

Microscopy – localization of GFP-ComEA and GFP-ComFA

Fluorescence microscopy was performed as previously described (162, 163). Exposure times were typically 500 ms for GFP-ComEA and 250 ms for GFP-ComFA. Membranes of *gfp-comEA* cells were stained with TMA-DPH (Molecular Probes), at a final concentration of 0.01 mM, and imaged with exposure times of 200 ms. Cell bodies of *gfp-comFA* cells were imaged using phase-contrast microscopy (20 ms exposures). Fluorescence images were analyzed, adjusted, and cropped using Metamorph v 6.1 software (Molecular Devices).

Microscopy – post-processing of Alexa Fluor 488-maleimide labeled ComGC pili and Label IT-Cy5 labeled PCR product images

All epifluorescence microscopy images presented in Fig. 12 through Fig. 14 were deconvoluted prior to publication. First, point-spread-functions (PSFs) were computationally estimated for Alexa Fluor 488 and Cy5 signals using the PSF Generator plugin (EPFL Biomedical Imaging Group) on Fiji (NIH) (164). For Alexa Fluor 488 signal captured on the Zeiss Axio Observer.Z1 microscope configured as described above, the following parameters were entered into the PSF Generator: Optical model = Born and Wolf 3D Optical Model ; Refractive index immersion = 1.518 (corresponding to Immersol 518F) ; Accuracy computation = Best ; Wavelength = 516 nm (corresponding to Alexa Fluor 488 emission maximum) ; NA (numerical aperture) = 1.4 ; Pixelsize XY = 65 nm ; Z-step = 250 nm ; Size XYZ – X = 256, Y = 256, Z = 65 ; Display = Linear, 16-bits, Fire. The Z-stack composing the PSF was Z-projected to a single 256 x 256 pixel image by averaging the pixel intensities of each individual pixel in the 256 x 256 pixel array across the 65 Z-slices of the Z-stack, resulting in a 2D PSF. For Alexa Fluor 488 signal captured on the Zeiss Axio Imager.Z2 microscope configured as described above, the same procedure was performed to estimate the PSF, only changing Pixelsize XY to 110 nm. For Cy5 signal captured

on the Zeiss Axio Imager.Z2 microscope configured as described above, all parameters were kept constant to estimate the PSF, only changing Wavelength = 666 nm (corresponding to Cy5 emission maximum).

Epifluorescence microscopy images were deconvolved using the DeconvolutionLab2 plugin (EPFL Biomedical Imaging Group) for Fiji (NIH). Individual epifluorescence microscopy images were opened in Fiji, and the DeconvolutionLab2 plugin was run. The PSF generated as above, corresponding to a particular epifluorescence signal, was used to deconvolve that signal with DeconvolutionLab2 (Algorithm = Richardson-Lucy, 100 iterations).

Quantification of pili produced in comGC co-expression strains

Images of Alexa Fluor 488-maleimide labeled cells from each co-expression strain, acquired as described previously, were opened in the Fiji image processing package (ImageJ2, NIH). The plugin ObjectJ was first used to identify 300 individual cells of each strain from phase-contrast images. Narrow filaments of length greater than 0.5 µm that were directly connected to the cell bodies of these 300 cells, which were surmised to be ComGC pili, were identified and counted in the green epifluorescence channel. The extracellular space immediately adjacent to the 300 cell bodies was scanned in the green epifluorescence channel for foci, which were likely sheared ComGC pili, and foci within 0.5 µm of the cell body were counted. These epifluorescence data were graphed using Microsoft Excel.

Measurement of ComGC pili lengths

Images of Alexa Fluor 488-maleimide labeled cells of the *comGC*^{E56C} and *comGC*^{S65C} co-expression strains, acquired as described previously, were opened in Fiji (NIH). The images were scaled up 5-fold using bilinear interpolation, and the phase-contrast and green epifluorescence

image channels were split for each pilus-producing cell. Each channel was converted into a binary image using Fiji's default thresholding parameters, and then outlines of the signal present in each channel were produced. The outline view of the phase-contrast channel demarcated the cell boundary and extracellular space, while the green epifluorescence channel outline divided a pilus from the extracellular space. These outlines were merged together to simultaneously display both the cell boundary and pilus boundary. Pilus length was manually measured by placing the start of a segmented line on the cell boundary, approximately where the medial axis of the pilus would cross the cell boundary, and creating a line that roughly followed the pilus medial axis to the extreme tip of the pilus.

ComGC pilus retraction rate measurements

Time lapse microscopy images of Alexa Fluor 488-maleimide labeled *comGC*^{S65C} cells, acquired as described previously, were visually scanned for potential ComGC pilus retraction events using Fiji (NIH). Once identified, cells with retracting pili were isolated, and pilus length was measured for each time point of the time lapse as described previously. Frame-to-frame retraction rates were calculated by dividing the change in pilus length between frames by the time interval of the time lapse. These data were collected for ten individual pilus retraction events, which included 44 instances of frame-to-frame retraction. Microsoft Excel was used to perform the t-test and ANOVA (using the Real Statistics Resource Pack release 8.6.3) referenced in the results section.

Localization analysis of fluorescent proteins (ComGC, ComEA, ComFA) and PCR product

Images acquired as previously described for the Alexa Fluor 488-maleimide labeled *comGC*^{S65C} co-expression strain. *gfp-comEA*, and *gfp-comFA* strains were opened in Fiji (NIH). For labeled ComGC^{S65C} pili, individual pilus-producing cells were isolated and identified in the phase-

contrast channel, and a 4x4 pixel white square was added to the green epifluorescence channel where the pilus medial axis would cross the cell boundary to mark the base of pilus production. Because maleimide labeling also resulted in cell membrane staining, cell bodies were checked against the green fluorescence channel to identify cell septa not visible in phase-contrast images. If a septum was identified, a 1-pixel wide white line was added onto the phase-contrast image across the cell body to allow MicrobeJ to detect multiple cells in the image; this process was repeated for 232 individual pilus production events. MicrobeJ (version 5.11x) was used to define cell bodies and medial axes from phase-contrast images, and pili bases were identified from the green epifluorescence images by adjusting the plugin's sensitivity parameters (tolerance and Z-score) until only the added white square was detected as a fluorescent focus (165). The identified cells and pili boundaries were associated with one another, and a heatmap of pilus production localization on a size-normalized rod-shaped cell was made within the "Heatmap(s)" tab of the results window. This heatmap was scaled up 10-fold using bilinear interpolation to create the final version included in Fig. 15.

For localization of GFP-ComEA foci, cell bodies were identified by TMA-DPH epifluorescence signal, and the outline of each cell was filled in using a white rounded rectangle with a 6-pixel wide black border to enhance MicrobeJ detection of cells. GFP-ComEA foci were identified in the green epifluorescence channel and marked with a 4x4 pixel white square to enhance contrast. Microbe J (version 5.11x) was used as described above to generate a heatmap of GFP-ComEA localization. For localization of GFP-ComFA foci, cell bodies of isolated single cells were identified using phase-contrast images, and GFP-ComFA foci were identified in the green epifluorescence channel. A localization heatmap was generated as described previously. For localization of Cy5-labeled PCR product binding to Alexa Fluor 488-maleimide labeled cells,

time course images acquired as described previously were opened in Fiji (NIH). a binding event was defined as the co-localization of a Cy5 focus with either the cell body or a labeled ComGC filament for at least 3 consecutive frames (≥ 15 s of co-localization). Cell bodies were identified as described previously, and Cy5-labeled PCR product foci were identified in the red epifluorescence channel and contrast-enhanced with a 4x4 white square as described previously. A localization heatmap was then generated as described previously.

Acknowledgments

The authors thank Daniel B. Kearns for supplying us with pKRH83 (*lacA::P_{comG}-comGC*) used in the mutagenesis and construction of *$comGC^{Cys}$* co-expression strains, Thorsten Mascher for providing us with the sequences for pBS0E and pBS0E-*P_{xyl}* which was used in the construction of the *P_{xyl}-rbs-comK* allele for natural competence induction, Jonathan Lombardino for constructing the violin plots found in Fig. 12C, and the entire Burton laboratory for thoughtful criticism, discussion, and comments. This work was supported by the Rita Allen Foundation Milton E. Cassel Award. JZ was supported by an NIH T32 training grant (GM07215).

Chapter 4: Discussion

Jason D. Zuke wrote this chapter with the editing and advice of Briana. M. Burton. Jonathan Lombardino and Tanya G Falbel assisted with editing of this chapter.

Overall significance of this work

Prior to this thesis, our knowledge of the mechanistic basis for DNA transport across the outer membrane/cell wall came primarily from a single Gram-positive and Gram-negative organism, *S. pneumoniae* and *V. cholerae*, respectively (96, 97). While these seminal studies greatly enhanced our understanding of the process, focusing only on these two model organisms may have belied the diversity of competence pilus systems across bacteria. *B. subtilis*, the Gram-positive model organism that was ironically the first Gram-positive observed producing oligomeric pilin homologue complexes during natural competence, has now joined the fray (80). Unlike *S. pneumoniae* and *V. cholerae*, *B. subtilis* forms natural competence pili that are likely stabilized by inter-pilin disulfide bonds, and both the average length and retraction rates of these disulfide bonded pili are significantly shorter/slower than their *S. pneumoniae* and *V. cholerae* counterparts (see Ch. 3). Although these differences are striking, the overall mechanism of DNA transport across the cell wall appears consistent across all three organisms. Each species produces pili during natural competence, these pili bind to DNA, and then retraction occurs to pull DNA across the outer cell periphery (see Ch. 1 – Fig. 7 and Ch. 3). Through this thesis, I have confirmed that natural competence pili can vary substantially between different species, but the general mechanism of DNA translocation across the outer cell periphery remains conserved.

Natural competence pilus length control and retraction rates

As noted above and in Ch. 3, the average (mean) length of natural competence pili varies significantly between *B. subtilis* (0.33 µm) and *S. pneumoniae* (0.5 µm)*/V. cholerae* (1 µm). The most salient difference between the ComG system in the Gram-positives and the more typical T4P system in *V. cholerae* is the presence of a single ATPase in ComG but multiple dedicated extension/retraction ATPases in the T4P system (76, 96, 97). Having two pools of ATPases with specialized functions in *V. cholerae* may make both proteins more efficient, versus a single ATPase pool that might be used for both extension and retraction in the ComG system. If this was true, then one might expect *V. cholerae* would polymerize a greater number of pilin subunits in a given unit of time than the Gram-positives, producing longer pili on average prior to retraction initiation. Alternatively, the ComG ATPases may simply be inherently less efficient than the T4P extension ATPases, which would produce the same results.

The more modest difference between *B. subtilis* and *S. pneumoniae* pilus lengths could be due to a number of different factors. Although both species produce ComGA, these homologues are only 55% similar in amino acid composition (34% identical), and as such could have differing pilus extension efficiencies based on those changes ($ComGA_{Bs}$: Uniprot P25953; $ComGA_{Sp}$: Uniprot B2IMM8). The other obvious difference between the two systems is the presence of the disulfide bonds in $ComGC_{Bs}$ but not in $ComGC_{Sp}$. *B. subtilis* effectively has two extra steps that presumably occur during the addition of each monomer, that being the breakage of the intramolecular disulfide in the ComGC monomer to be added, then the formation of an intermolecular disulfide bond between the newly incorporated monomer and the next monomer in the filament. Depending on the reaction kinetics of these steps, this may slow the incorporation of pilin subunits into the pilus, resulting in shorter *B. subtills* pili.

Virtually the same arguments can be made concerning the differences in natural competence pilus median retraction rate across these various organisms (*B. subtilis* (12 nm/sec); *S. pneumoniae* (80 nm/sec); *V. cholerae* (100 nm/sec)). Differences in ATPase efficiencies or the kinetics of pilus disassembly could result in retraction rate differences. In addition, it is possible that each pilus has a different propensity for disassembly, which could inherently influence the overall retraction rates observed.

DNA binding to natural competence pili

Although maleimide labeling has drastically improved our understanding of DNA translocation across the outer cell periphery, how DNA binds to natural competence pili remains poorly understood. Is there a conserved mechanism of DNA binding to competence pili, and how is this binding achieved? The necessity of a specific minor pilin (ComP) for DNA binding to *Neisseria* T4P both *in vivo* and *in vitro*, the existence of *V. cholerae* minor pilin point mutants that significantly reduce DNA binding and natural transformation while still allowing for wild type piliation, and the known localization of minor pilins to the tips of T4P (where DNA appears to bind in maleimide labeling experiments) all point to a significant contribution of minor pilins in the binding of T4P to DNA (92, 96, 110, 114). Binding DNA at the pilus tip minimizes the overall width of a pilus-DNA complex, which could reduce steric hindrance when DNA ultimately passes through the outer membrane secretin and/or the pore formed in the cell wall from pilus extrusion, making DNA entry more feasible. The lack of homology between known DNA-binding minor pilins, including a recently characterized minor pilin (ComZ) from *Thermus thermophilus*, and the known preference for *Neisseria* ComP to bind DNA sequence-specifically, suggests that the evolution of minor pilin DNA binding properties may have occurred independently across numerous clades (92, 146). The underlying mechanism of DNA binding might, therefore, be variable across natural competence pilus systems.(92, 146)

The necessity of the ComGC$_{Bs}$ disulfide bond

The essential disulfide bond in *B. subtilis* ComGC is a curious observation, in light of *S. pneumoniae* ComGC having no cysteines yet functioning adequately for natural transformation (80, 97). The question thus arises: what function does the ComGC disulfide bond have in the context of *B. subtilis* pilus filaments? Each ComGC monomer in the filament being disulfide linked would undoubtedly increase the stability of these filaments (80)due to the covalent nature of these bonds.

Perhaps the environments *B. subtilis* most likely evolved in, the soil and regions adjacent to plant roots, necessitated hardier filaments than those produced in *S. pneumoniae* (found in respiratory tracts and nasal cavities) and *V. cholerae* (found in aquatic environments associated with chitinous surfaces) (166–168). Thinking about these environments, the amount of bulk particulate matter in the *B. subtilis* natural environment would probably be much greater than in the environments of the other two organisms discussed. This might expose elongating pilus filaments to a large amount of physical resistance. Such resistance could induce extreme bending within the filaments as the pilus tip is diverted away from its original path, resulting in breaks along the filament. These breaks would be particularly deleterious if the pilus tip is involved in DNA binding (as discussed above) since any break would remove the tip from the rest of the filament, making the apparatus useless for DNA uptake. Evolving a stronger filament through disulfide bonding could be greatly beneficial in this scenario.

Another possibility is that, in reality, ComGC pili are not actually stabilized by intermolecular disulfide bonds. The only evidence supporting intermolecular disulfide bond formation is an observation that, by Western blot of protein samples separated by SDS-PAGE, ComGC appears oligomerized in non-reduced samples but monomeric when reducing agents are added (80). This

inference is problematic for two main reasons. First, strong protein denaturing agents such as 8M guanidine can fail to denature pili that are not disulfide linked, so the presence of ComGC oligomers in non-reducing SDS-PAGE gels does not inherently imply the presence of disulfide bonds between ComGC monomers (169). Secondly, the observation of ComGC pilus denaturation upon addition of reducing agents could plausibly occur if ComGC monomers within pili exhibited the same intramolecular disulfide bonds observed in non-polymerized monomers (105). If the polymerized ComGC intramolecular disulfide bonds were accessible to reducing agents, then exposing pili to these reducing agents could cause breakage of the disulfide bonds and structural instability of pilin monomers within filaments, thereby leading to collapse of ComGC filaments but not necessarily the degradation of monomers. The successful generation of experimentally derived ComGC pilus structural models in the future will be indispensable for conclusively determining if intermolecular disulfide bonds form during ComGC pilus production.

Future directions

Structural biology to elucidate natural competence pilus – DNA binding interactions

Recent advancements in both pilus purification and cryogenic electron microscopy techniques have set the stage to understand the competence pilus – DNA binding interaction more comprehensively. Morais et al. demonstrated that pili from *L. rhamnosus* could be efficiently purified in high concentration by a simple protocol consisting of cell wall digestion in an osmotically protective buffer to liberate pili without significant cell lysis, followed by two chromatography steps to remove impurities (170). The underlying principles at work in this purification should be widely applicable to other pilus systems. Therefore, intact competence pili from a diverse array of competent bacteria should now be readily purifiable for downstream *in vitro* DNA interaction studies. This is especially important for any pilus system that depends on the quaternary structure of the pilus or a pilin complex for DNA binding, as purification of individual pilus components would be insufficient for DNA binding studies in these cases.

Cryogenic electron microscopy (cryoEM) can be leveraged to deduce the key molecular interactions stabilizing competence pilus – DNA binding. Molecular structure determination via cryoEM involves the rapid freezing of samples in amorphous ice (vitrification), collecting anywhere from $10^4 – 10^6$ images of the relevant molecule(s) in the sample in various orientations using electron microscopy, and then computationally reconstructing a 3D model of the molecule(s) from the electron micrographs (reviewed in (171, 172)). Improvements in cryoEM technology (mainly in electron detection and image processing) have allowed for electron density map resolutions < 4 Å, which are generally sufficient for generating atomic models (173).

Notably, cryoEM does not require the crystallization of target molecules or any fixation procedure prior to vitrification. Therefore, any purified competence pilus can be incubated with substrate DNA *in vitro* to allow binding to occur, and the sample can then be directly vitrified and subjected to electron microscopy for structural determination. As long as a high enough resolution is achieved, an atomic model should be possible to build that depicts the critical binding interactions between pilus and bound DNA. This process can be applied to numerous bacterial competence pilus systems, and a comparative analysis of the pilus – DNA binding interactions could be performed to assess the diversity of these interactions. Structural determination via cryoEM was recently used to determine how MutS dimers bind and respond to DNA mismatches to subsequently initiate the mismatch repair pathway, which supports the utility of cryoEM for studying interactions within nucleoproteins (174).

Determining how ComG pili retract

There are two known pilus retraction mechanisms consistent with a single ATPase, as is the case in the ComG pilus systems. In the first, the lone ATPase is bifunctional, able to both extend and retract the pilus. The *C. crescentus* ATPase CpaF exhibits bifunctionality for extending and retracting tight adherence pili that are essential for surface colonization (85). Secondly, pilus retraction has been shown to be spontaneous. *V. cholerae* competence pili retract at greatly reduced rates in the absence of the retraction dedicated ATPase pair PilTU, and this spontaneous retraction appears to stem from the preference of the terminal pilin subunit to associate with the membrane over binding to the next pilin in the filament, which leads to dissociation of the filament over time in the absence of active extension (147).

Differentiating between these two mechanisms in the competent Gram-positives will be difficult. The elucidation of CpaF's bifunctionality was greatly facilitated by a phage whose infectivity is

dependent on pilus retraction, which allowed for the selection of piliated, yet retraction-deficient, cpaF mutants (85, 147). To our knowledge, no such phage has been characterized for either of the Gram-positive models, and the transformation frequencies of both models (~10^{-5} – 10^{-4} transformants/CFU) are far too low to rely on transformation of a counter-selectable marker to cull retractile comGA mutants from an otherwise piliated comGA mutant library. For the time being, this essential step in the Gram-positive transformation process remains poorly understood.

Assessing the necessity of ComGC disulfide bonds for B. subtilis pilus production

The $ComGC_{Bs}$ disulfide bond, which I've already discussed extensively, is a unique characteristic of the *B. subtilis* natural competence pilus. The necessity of the disulfide bond for $ComGC_{Bs}$ monomer stability is clear, however that does not imply that disulfide bonds (either intra- or intermolecular) are required for DNA import during *B. subtilis* natural competence in all circumstances (80). To test if ComGC pilus production is absolutely dependent on disulfide bond formation in *B. subtilis*, a relatively straightforward experiment would be to complement the Δ*comGC*$_{Bs}$ allele used in Ch. 2 with ectopically expressed *comGC*$_{Sp}$. Restoration of natural transformability in that strain would imply successful production of a $ComGC_{Sp}$ pilus in *B. subtilis*, providing direct evidence that disulfide bond formation is dispensable for pilus formation. Because of evolutionary divergence of the rest of the *comG*$_{Sp}$ operon, simply substituting *comGC*$_{Sp}$ may not be sufficient for pilus production. In that case, allelic replacement of the entire *comG*$_{Bs}$ operon with *comG*$_{Sp}$ would be more appropriate.

Appendix A: Heterologous production of the ComGC soluble domain for solution NMR

Jason D. Zuke conceptualized and wrote this chapter. Briana M. Burton assisted with conceptualization. Chad M. Rienstra assisted with conceptualization and provided access to necessary personnel and equipment. Katarzyna Gromek, Ronny Fredrick, Paulo Cobra, and Marco Tonelli all contributed to performing experiments for the completion of this chapter.

Introduction

The intra- to intermolecular disulfide bond transition that may occur upon ComGC oligomerization into a pilus is poorly understood (80, 105). In an effort to both better understand how such a transition might occur (if at all), as well as how ComGC pili are structured, I wanted to work towards reconstructing a ComGC pilus filament from structural data. The general plan was a cryoEM reconstruction of the filament by fitting a ComGC monomer structural model into the electron density of a ComGC filament, in the same vein as other pilus cryoEM reconstructions (107). Because there was no available structural data for *B. subtilis* ComGC, I needed to first work on solving the structure of monomeric ComGC for model building. Owing to the small size of the monomer (~ 10 kDa), solution NMR was a good option. The following sections will detail my (and my collaborators) attempts at heterologously producing the soluble portion of ComGC in *E. coli* and solving the structure via solution NMR.

Results

Designing the heterologous ComGC soluble domain expression vector

To improve the chances of high yield heterologous expression of *B. subtilis comGC* in *E. coli*, I designed an expression vector containing a *comGC* allele that will express an N-terminally truncated variant of ComGC, under control of the T7 promoter. This variant, lacking residues 1-30, removes the signal sequence targeting ComGC to the cytoplasmic membrane and the transmembrane portion of the N-terminal alpha helix. Cytoplasmic production of ComGC$^{\Delta 1\text{-}30}$ should then be possible, greatly facilitating higher yields as seen with pilins from different systems (106, 175). For downstream purification purposes, the *comGC$^{\Delta 1\text{-}30}$* allele was fused with the gene encoding the SUMO domain tag, which itself contained a 6X-His tag at the 5' end. IMAC can thus be used to capture the SUMO-ComGC$^{\Delta 1\text{-}30}$, and ComGC$^{\Delta 1\text{-}30}$ released from the IMAC column by incubation with SUMO protease.

Expression and purification of ^{15}N-labeled ComGC soluble domain

The *6XHis-SUMO-comGC$^{\Delta 1\text{-}30}$* expression vector was transformed into *E. coli* BL21(DE3)pLysS for heterologous expression. This resulting strain was cultured in 1 L total of ^{15}N-labeling M9 minimal medium and induced with IPTG. Cells were lysed, clarified, and the lysate was run through an IMAC column to bind the resulting 6XHis-SUMO-ComGC$^{\Delta 1\text{-}30}$. An imidazole buffer was added to elute the protein from the column, the eluent dialyzed against SUMO protease cleavage buffer, then was treated with SUMO protease to separate the 6XHis-SUMO tag from ComGC$^{\Delta 1\text{-}30}$. This mixture was once again run through an IMAC column to purify ComGC$^{\Delta 1\text{-}30}$ in the flowthrough. This procedure resulted in a relatively pure, highly concentrated sample of ComGC$^{\Delta 1\text{-}30}$ (Sup. Fig. 6). As a final polishing step, the sample was subjected to gel filtration chromatography. Fractions containing ComGC$^{\Delta 1\text{-}30}$ were pooled and concentrated using a

centrifugal concentrator, then dialyzed against NMR buffer. The resulting sample contained highly pure, highly concentrated ComGC$^{\Delta 1\text{-}30}$ for NMR (Sup. Fig. 7).

Solution NMR of ^{15}N-labeled ComGC soluble domain

The ^{15}N-labeled ComGC$^{\Delta 1\text{-}30}$ sample was subjected to both ^{15}N HSQC and ^{1}H-^{15}N heteronuclear NOE experiments. Both NMR experiments were consistent with ComGC$^{\Delta 1\text{-}30}$ containing a small portion of residues within structured domains, but also a significant number of residues being in domains that were either flexible or unstructured (Sup. Fig. 8). The poor chemical shift dispersion observed in the ^{15}N HSQC argued against a well ordered ComGC$^{\Delta 1\text{-}30}$ sample (Sup. Fig. 8A). Similarly, the negative peaks associated with the majority of chemical shifts in the ^{1}H-^{15}N heteronuclear NOE spectrum are indicative of significant disorder or flexibility within ComGC$^{\Delta 1\text{-}30}$ (Sup. Fig. 8B). Because of this lack of structure and/or flexibility, the chances of obtaining a high-resolution structural model of ComGC based on these data were low.

Discussion

While we were ultimately unable to produce a ComGC soluble domain prep amenable to structural determination by solution NMR, many positives came out of these experiments. First, a SUMO-tagged ComGC soluble domain fusion can be produced to a high concentration in the heterologous *E. coli* protein production strain employed here (Sup. Fig. 6). Secondly, the purification workflow presented in this study was successful, allowing for extremely pure and highly concentrated ComGC soluble domain samples (Sup. Fig. 7). Taken together, these data confirm that future attempts at purifying ComGC soluble domain will be worthwhile to pursue.

Both the ^{15}N HSQC and ^{1}H-^{15}N heteronuclear NOE experiments suggested significant disorder or flexibility within our ComGC soluble domain prep (Sup. Fig. 8). Arriving at a high-resolution structural model with such constraints is not realistic, so we opted to forgo the more expensive ^{13}C/^{15}N ComGC soluble domain prep until we could address this problem. The most promising avenue for a more ordered ComGC soluble domain structure is the purification of the protein in an oxidizing environment, rather than the reducing environment found across the entire set of previous purifications and NMR experiments. ComGC is stabilized by an intramolecular disulfide bond when produced in *B. subtilis*, so this may be a key factor in structuring the protein (105). Heterologously producing the protein in the *E. coli* SHuffle strain, which maintains an oxidizing environment in the cytosol, coupled with purification steps lacking reducing agents could drastically improve the folding and rigidity of the resulting ComGC soluble domain in future experiments (176).

Materials and methods

Bacterial strains and strain construction

The only strain used in this study was a derivative of *E. coli* BL21(DE3)pLysS bearing an IPTG-inducible expression plasmid for *6XHis-SUMO-comGC$^{\Delta 1\text{-}30}$* (pJZ085). pJZ085 was constructed commercially by GenScript gene synthesis and subcloning services, then transformed into chemically competent *E. coli* BL21(DE3)pLysS by heat shock (177).

Media and growth conditions

E. coli bearing pJZ085 was cultured in Lennox's Lysogeny Broth + 0.2% (w/v) glucose (LB-Lennox + glucose) at either 25 °C or 37 °C for general propagation, and ^{15}N-labeling M9 Minimal Medium (standard M9 minimal medium supplemented with ^{15}N-NH_4Cl) at either 20 °C or 37 °C for production of ^{15}N-labeled protein (178). Cells were plated onto Miller's Lysogeny Broth + 0.2% (w/v) glucose (LB-Miller + glucose) 1.5% (w/v) agar plates. All media was supplemented with 35 μg/mL chloramphenicol and 50 μg/mL kanamycin, except for the media used prior to plating pJZ085 transformants, which was supplemented with only 35 μg/mL chloramphenicol as pJZ085 had not yet been transformed.

Purification of ComGC$^{\Delta 1\text{-}30}$ (ComGC soluble domain)

A freshly transformed single colony of *E. coli* BL21(DE3)pLysS bearing pJZ085 was inoculated into 1 mL LB-Lennox + glucose and cultured at 37 °C with 250 rpm shaking until turbidity could be observed. This entire culture was transferred into 25 mL of the same medium and incubated for 16 hr at 25 °C with 250 rpm shaking. This entire culture was added to 1 L ^{15}N-labeling M9 Minimal Medium, split into 2 x 500 mL cultures, then cultured at 37 °C until the OD_{600} = 1.0. At this point, the temperature was decreased to 20 °C and incubated for an additional 30 min. 0.4

mM IPTG was added to the cultures, and incubation continued for an additional 16 hr to produce the final culture for processing.

The cultures were centrifuged at 4,300 x RCF for 20 minutes to pellet the cells, and the supernatant decanted completely. The pellets were resuspended in a total of 100 mL of IMAC binding buffer (50 mM Tris pH 8.5, 150 mM NaCl, 5 mM imidazole), and then lysed by sonication. The raw lysate was passed through a 1.2 μm cellulose acetate filter to remove large particulates, and then added to a pre-prepared IMAC purification column and allowed to flow through by gravity flow. The ComGC soluble domain fusion was eluted from the column by adding 20 mL IMAC elution buffer (50 mM Tris pH 8.5, 150 mM NaCl, 500 mM imidazole, 1 mM beta-mercaptoethanol) and collecting the resulting eluent. The flowthrough was dialyzed against SUMO protease cleavage buffer (50 mM Tris pH 8.5, 150 mM NaCl, 5 mM imidazole, 1 mM beta-mercaptoethanol), then treated with SUMO protease to remove the 6XHis-SUMO tag from ComGC$^{\Delta 1\text{-}30}$. The sample was then run through an IMAC column to capture the cleaved 6XHis-SUMO and leave the ComGC$^{\Delta 1\text{-}30}$ in the flowthrough.

The sample was then concentrated to approximately 1 mL using a regenerated cellulose centrifugal concentrator with a 3 kDa molecular weight cut off. As a final polishing step, the concentrated sample was run through Sephacryl S-100 16/60 gel filtration resin. The fractions containing ComGC$^{\Delta 1\text{-}30}$ were determined by SDS-PAGE and Coomassie staining, then these fractions were pooled and dialyzed against NMR buffer (50 mM HEPES pH 6.02. 150 mM NaCl, 0.5 mM TCEP). This dialyzed sample was then concentrated to ~ 500 uL using the centrifugal concentrator mentioned above, and the concentration of ComGC$^{\Delta 1\text{-}30}$ was determined.

Acknowledgments

The authors would like to thank Dr. Chad Rienstra and NMRFAM for consulting with us on this study and allowing us access to an NMR spectrometer with technical assistance. We thank Dr. Ronny Fredrick for helping design the ComGC soluble domain fusion construct, producing the *E. coli* expression clones, as well as designing and running the ComGC soluble domain purification strategy. We appreciate the work of Dr. Paulo Cobra and Dr. Marco Tonelli for analyzing our ComGC soluble domain samples and helping us understand the significance of the resulting spectra. Finally, we thank Dr. Katarzyna Gromek for assisting with the purification of the ComGC soluble domain.

Figures

Figure 1

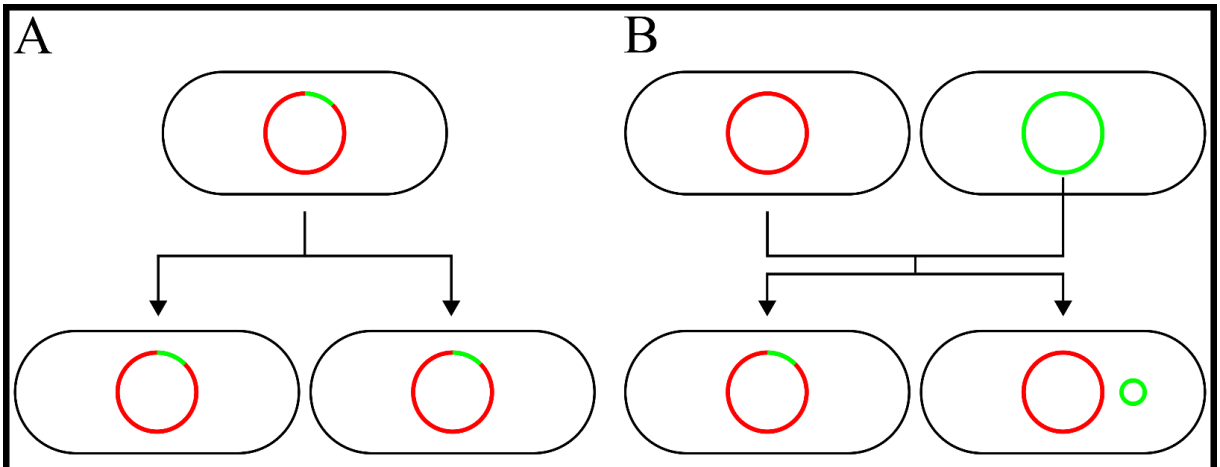


Fig. 1: Methods of bacterial gene transfer

(A) Diagram of bacterial vertical gene transfer. A mutation (green) occurs on the bacterial chromosome (red), and upon genome replication and cell division, that mutation is passed on to progeny.

(B) Diagram of bacterial horizontal gene transfer. A recipient cell (red chromosome) and a donor cell (green chromosome) from different cell lineages coexist in an environment. DNA from the donor cell is transferred into the recipient cell via phage transduction, conjugation, natural competence, or "non-canonical" mechanisms. Segments of the transferred DNA can be stably maintained in the recipient genome, either within a chromosome or extra-chromosomally, and propagate further by either vertical or horizontal gene transfer.

Figure 2

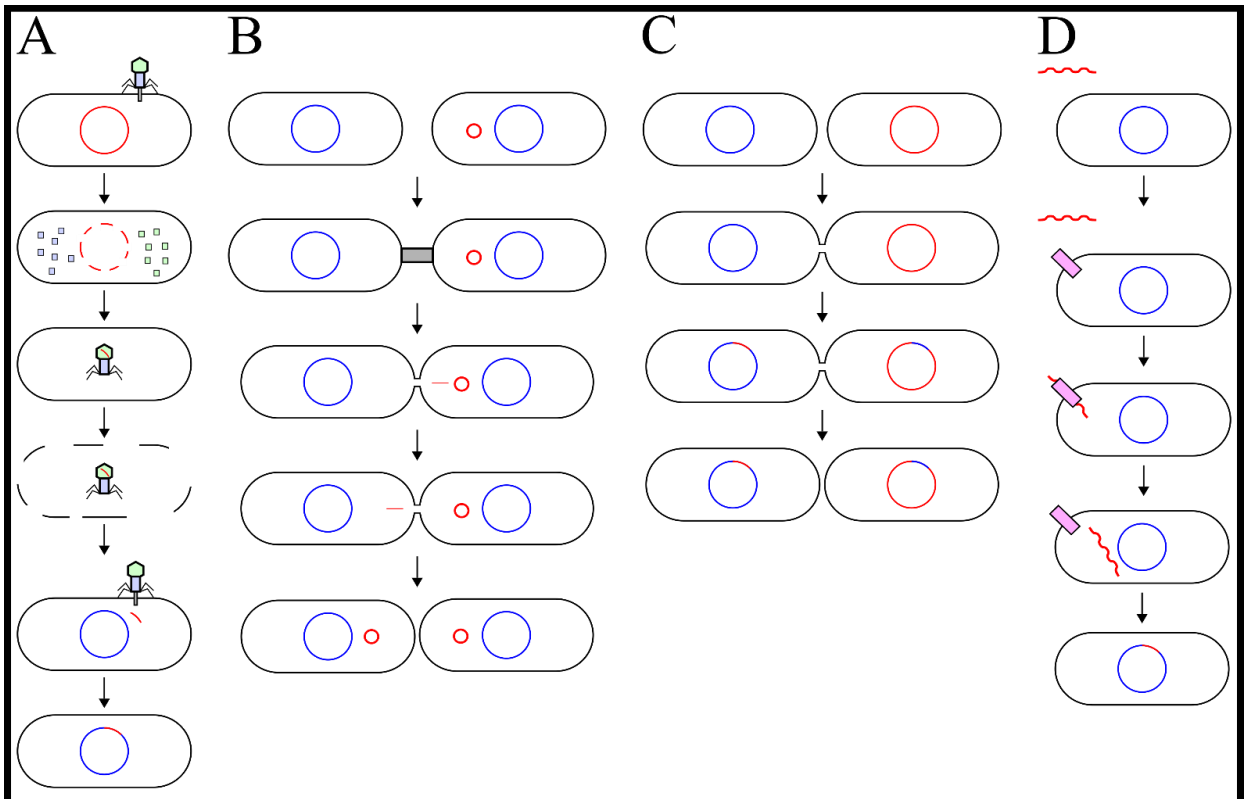


Fig. 2 Mechanisms of bacterial horizontal gene transfer

(A) Diagram of bacteriophage transduction. A phage (green/blue structure on the cell) infects a bacterial donor cell (red chromosome). During the phage lytic cycle, new phage structural proteins are produced, and the donor chromosome is fragmented due to stress from the phage infection. As the lytic cycle continues, phage virions are assembled, but occasionally a donor DNA fragment will be incorrectly packaged into virions. These fully infectious phage virions containing donor DNA can infect recipient cells (blue chromosome), introducing the donor DNA fragment into the recipient cell. The introduced donor DNA can then be integrated into the recipient chromosome, completing the gene transfer.

(B) Diagram of bacterial conjugation. Donor cells harboring a conjugative mobile genetic element (small red circle) can produce a type IV secretion system pilus that binds to recipient cells and brings cells in close proximity to one another by pilus retraction. A membrane fusion event occurs, joining the two cells together. The mobile genetic element will replicate via rolling circle replication, and the replicated ssDNA will be transported into the recipient cell through a secretion pore complex, where the complementary strand will be synthesized for stable incorporation of the conjugative element in the recipient cell.

(C) Diagram of bacterial nanotube transfer, a “non-conical” horizontal gene transfer mechanism. Cells in close proximity to one another, after long periods of growth, can become connected by presumably membranous “nanotubes”. Genetic transfer across these nanotubes has been documented, though the mechanistic steps are still under investigation.

(D) Diagram of bacterial natural transformation. Naturally competent cells can form a protein complex (purple rectangle) that imports extracellular DNA into the cytoplasm as single stranded DNA. This process is mediated by natural competence type IV pili which appear to be involved in transporting DNA across the outer membrane and/or cell wall, as well as numerous proteins which stabilize DNA at the cell surface and allow for DNA entry across a cytoplasmic membrane pore complex. At some point in the process, DNA is converted to ssDNA which can be incorporated into the recipient cell’s genome.

Figure 3

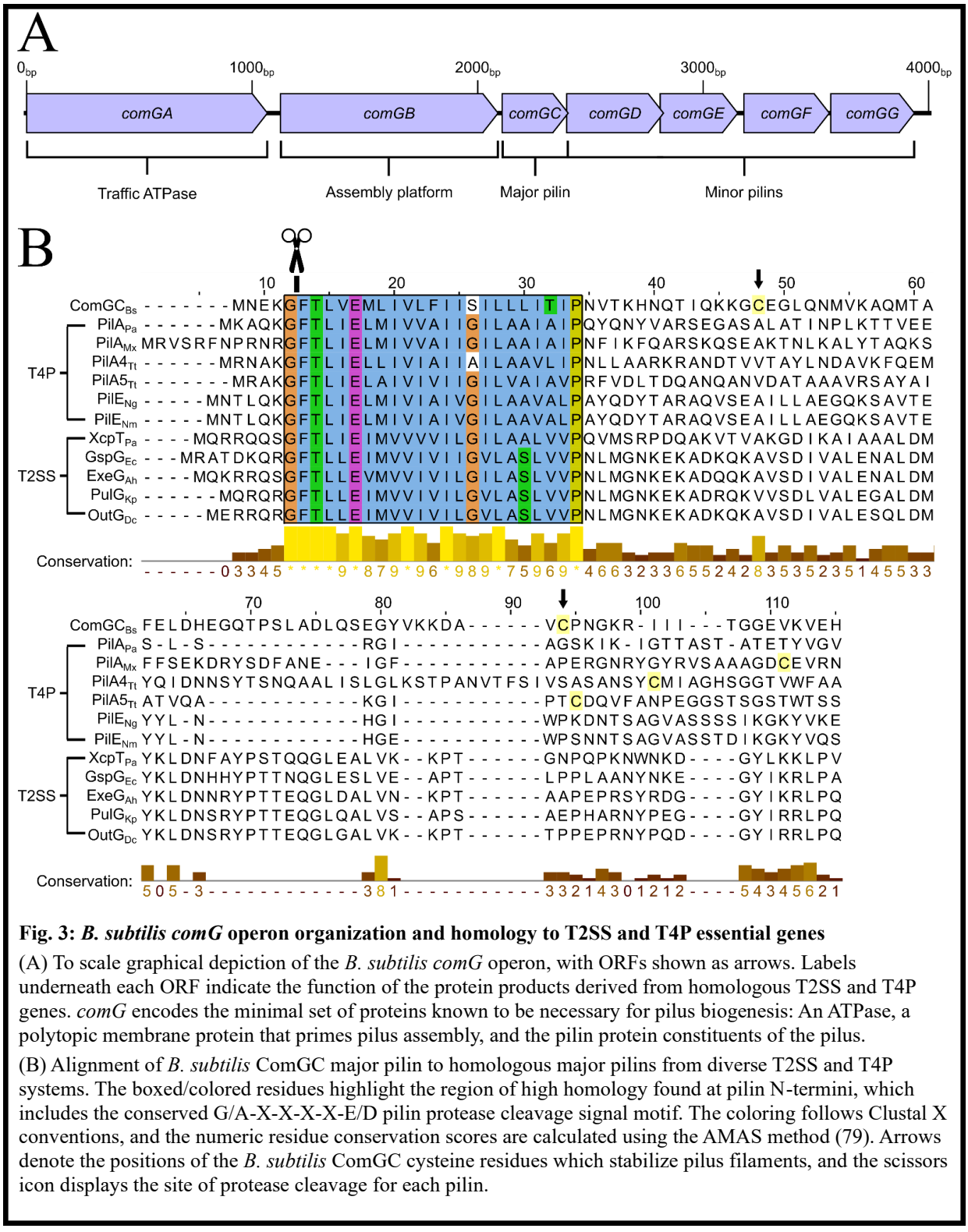


Fig. 3: *B. subtilis comG* operon organization and homology to T2SS and T4P essential genes

(A) To scale graphical depiction of the *B. subtilis comG* operon, with ORFs shown as arrows. Labels underneath each ORF indicate the function of the protein products derived from homologous T2SS and T4P genes. *comG* encodes the minimal set of proteins known to be necessary for pilus biogenesis: An ATPase, a polytopic membrane protein that primes pilus assembly, and the pilin protein constituents of the pilus.

(B) Alignment of *B. subtilis* ComGC major pilin to homologous major pilins from diverse T2SS and T4P systems. The boxed/colored residues highlight the region of high homology found at pilin N-termini, which includes the conserved G/A-X-X-X-X-E/D pilin protease cleavage signal motif. The coloring follows Clustal X conventions, and the numeric residue conservation scores are calculated using the AMAS method (79). Arrows denote the positions of the *B. subtilis* ComGC cysteine residues which stabilize pilus filaments, and the scissors icon displays the site of protease cleavage for each pilin.

Figure 4

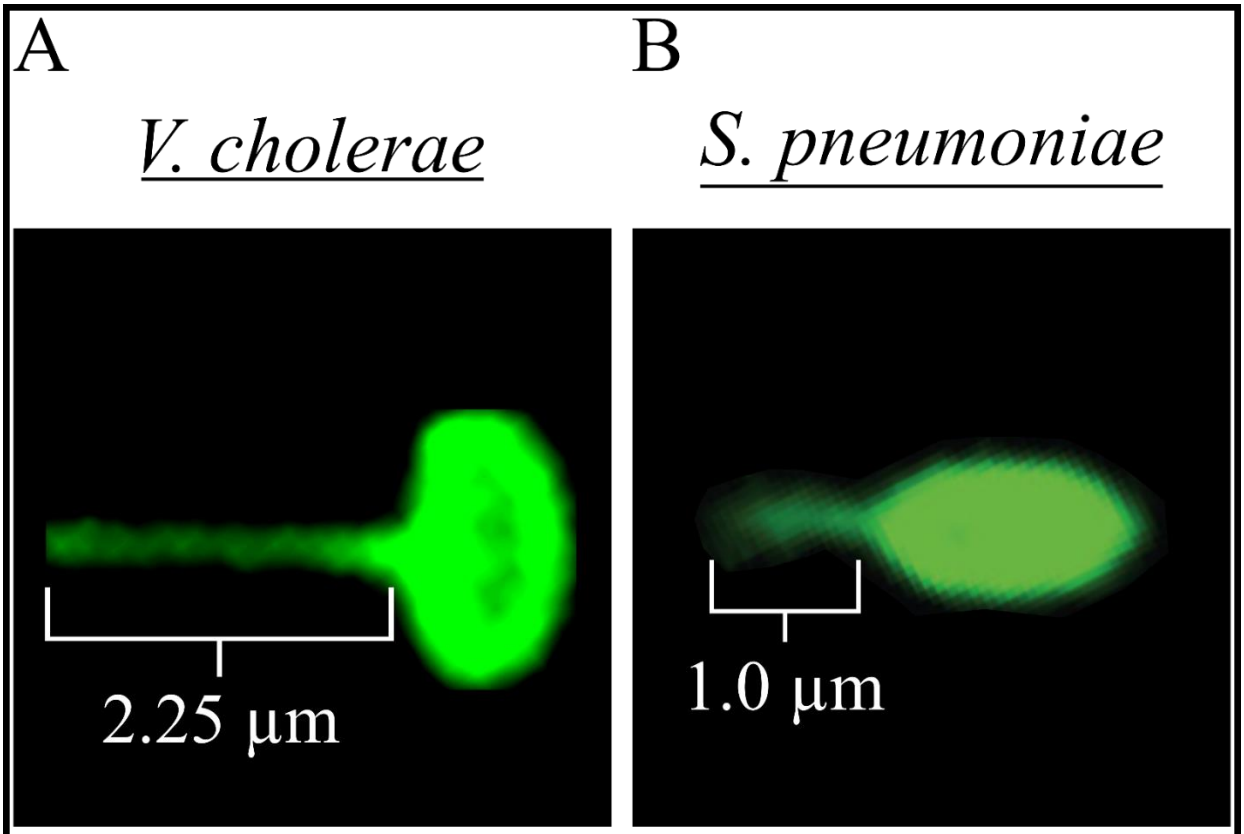


Fig. 4: Examples of natural competence pili fluorescently labeled via the maleimide labeling method

Widefield fluorescence micrographs of naturally competent cells producing pili (false-colored green) composed of cysteine substitution mutant pilin monomers, labeled by exposure to maleimide-conjugated Alexa Fluor 488. Labels indicate the length of each pilus shown. Organisms shown include (A) *V. cholerae* (96) and (B) *S. pneumoniae* (97). Figures adapted from the preceding references with permission.

Figure 5

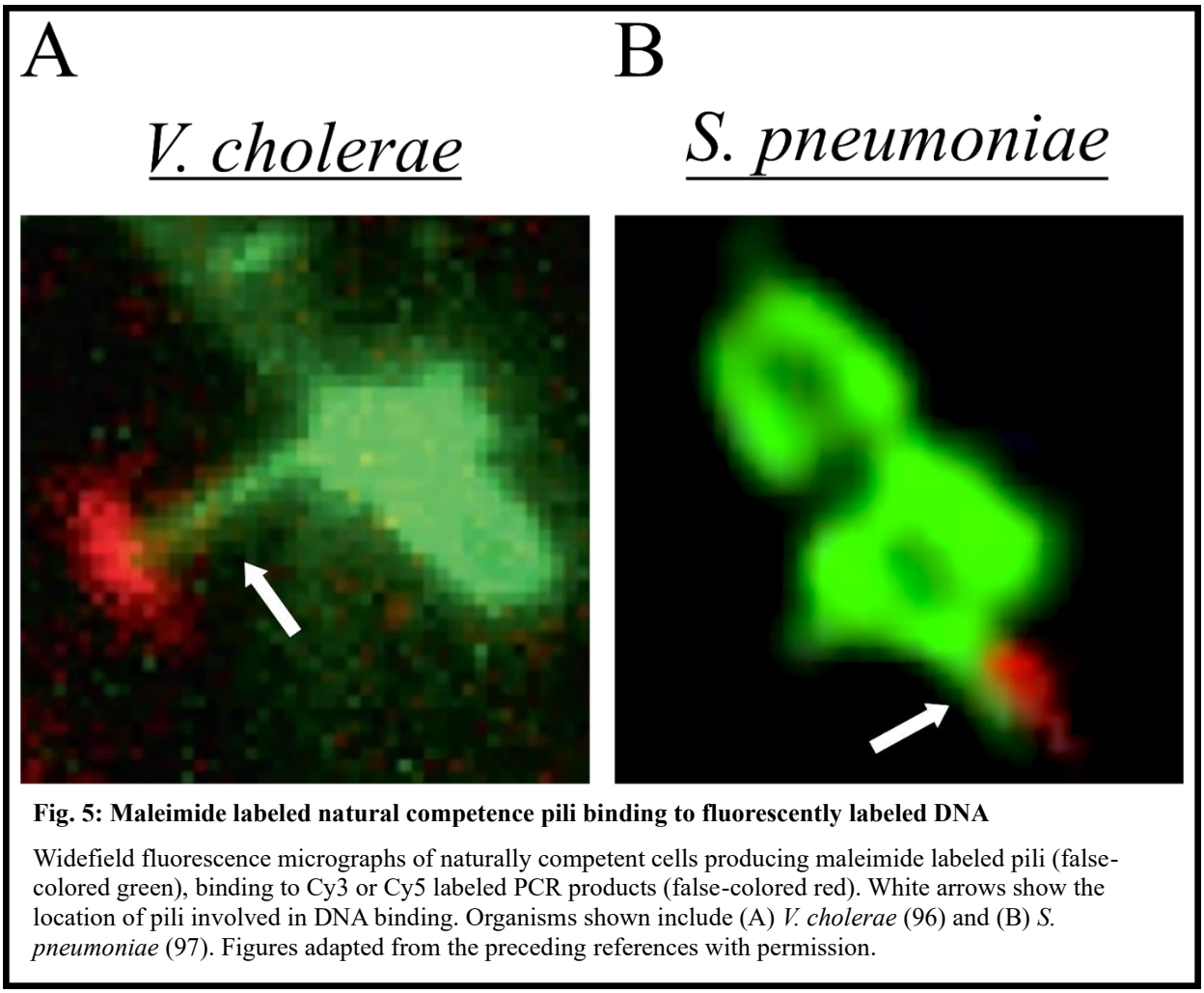


Fig. 5: Maleimide labeled natural competence pili binding to fluorescently labeled DNA

Widefield fluorescence micrographs of naturally competent cells producing maleimide labeled pili (false-colored green), binding to Cy3 or Cy5 labeled PCR products (false-colored red). White arrows show the location of pili involved in DNA binding. Organisms shown include (A) *V. cholerae* (96) and (B) *S. pneumoniae* (97). Figures adapted from the preceding references with permission.

Figure 6

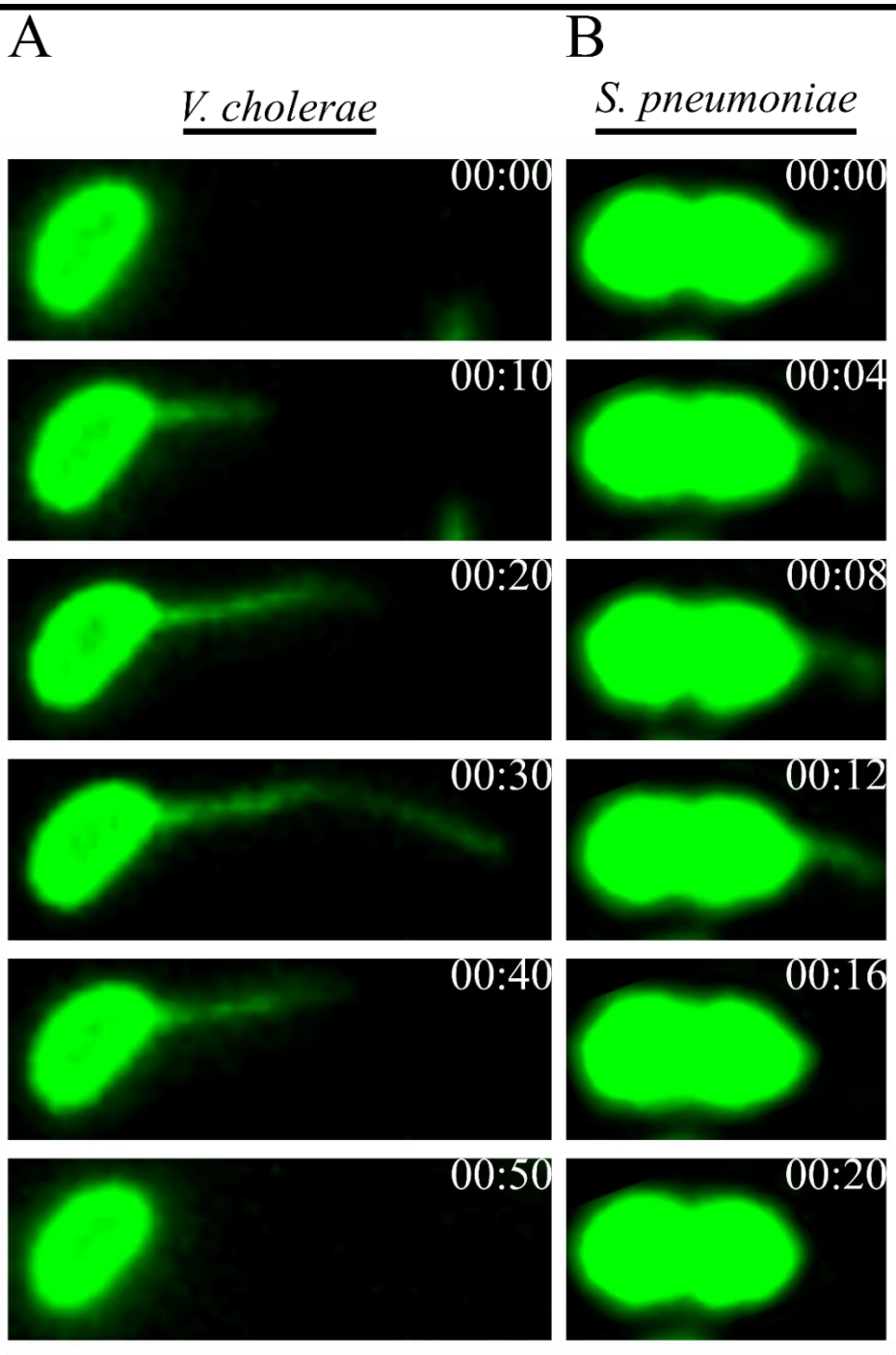


Fig. 6: Extension and retraction of natural competence pili

Widefield fluorescence micrographs of naturally competent cells extending and retracting maleimide labeled pili (false-colored green). The earliest timepoints are displayed in the top images, progressing through time to the latest timepoints displayed in the bottom images. Timescales vary for each organism and are displayed at the top-right corner of each image (min:sec label). Organisms shown include (A) *V. cholerae* (96) and (B) *S. pneumoniae* (97). Figures adapted from the preceding references with permission.

Figure 7

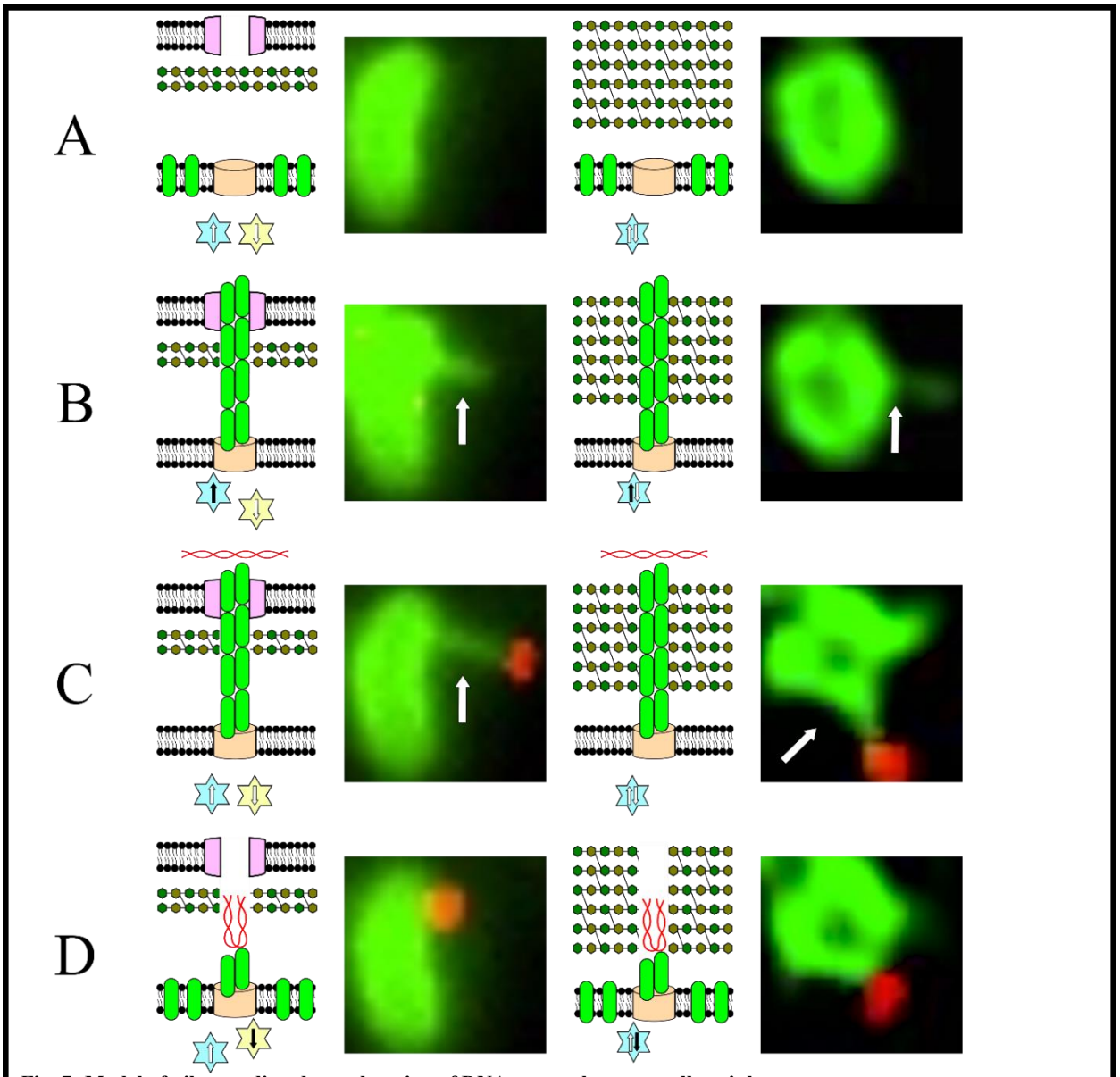


Fig. 7: Model of pilus-mediated translocation of DNA across the outer cell periphery

The far left and middle-right columns show a cartoon model of the outer cell periphery of a naturally competent Gram-negative and Gram-positive cell, respectively, during a given step in the DNA translocation process. The middle-left and far right columns show example widefield fluorescence micrographs of a green maleimide-labeled naturally competent Gram-negative (*V. cholerae*) and Gram-positive (*S. pneumoniae*) cell, respectively, during translocation of a red Cy3 labeled PCR product (96, 97). For the fluorescence micrographs, white arrows show the location of assembled pili. The cartoon model component identities are as follows: The purple rounded trapezoids in the outer membrane represent outer membrane secretins; the orange cytoplasmic membrane cylinder represents pilus assembly platform protein; the green rounded rectangles in the cytoplasmic membrane represent pilins; the blue/yellow six-sided stars represent pilus extension/retraction dedicated ATPases, respectively. For the Gram-positive cartoons, a single ATPase is present which presumably powers both extension and retraction. (A) Naturally competent cells produce proteins required for pilus biogenesis; (B) pilin monomers are oligomerized into pilus filaments through the concerted actions of platform proteins and extension ATPases, which breach into the extracellular space; (C) pili bind to free DNA molecules at their tips, most likely via minor pilins; (D) Retraction ATPases facilitate the disassembly of pilin monomers from the pilus base, which causes the pilus to retract and simultaneously pull bound DNA through the pores created in the outer cell periphery. Figures adapted from the preceding references with permission.

Figure 8

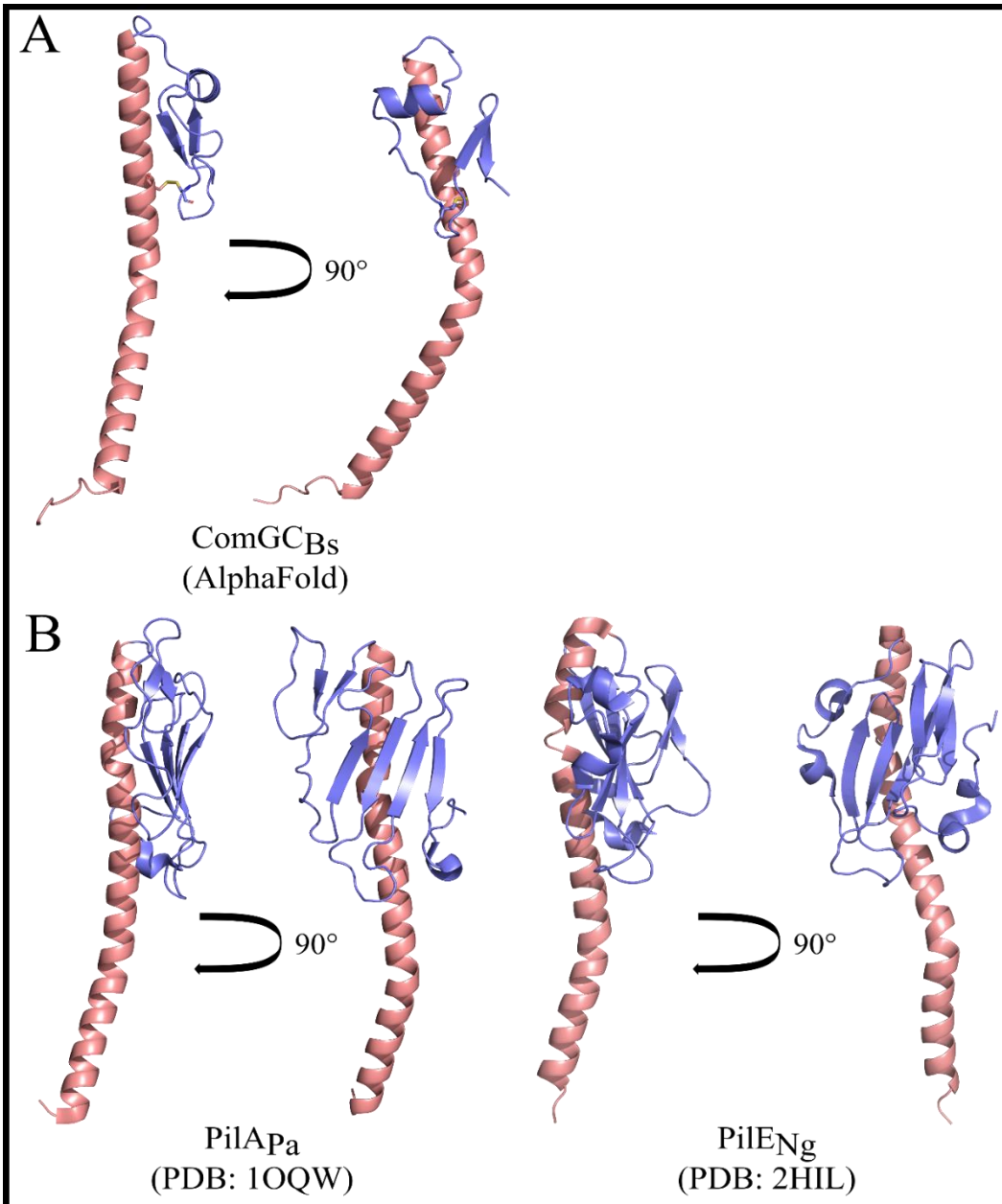


Fig. 8: *Bacillus subtilis* ComGC is likely to be structurally homologous to known pilins

(A) Two views of a cartoon of the highest ranked AlphaFold2 structure of monomeric *Bacillus subtilis* ComGC. An N-terminal, pilin-like extended alpha helix is colored in salmon red, while the globular C-terminal domain is colored in slate blue. The intramolecular disulfide bond between Cys-41 and Cys-81 is highlighted in yellow and also depicted as both cartoon and sticks.

(B) Two views of cartoons of two experimentally determined T4P major pilin structures from *Pseudomonas aeruginosa* (PilA$_{Pa}$; PDB: 1OQW) and *Neisseria gonorrhoeae* (PilE$_{Ng}$; PDB: 2HIL). The coloring scheme is the same described in (A).

Figure 9

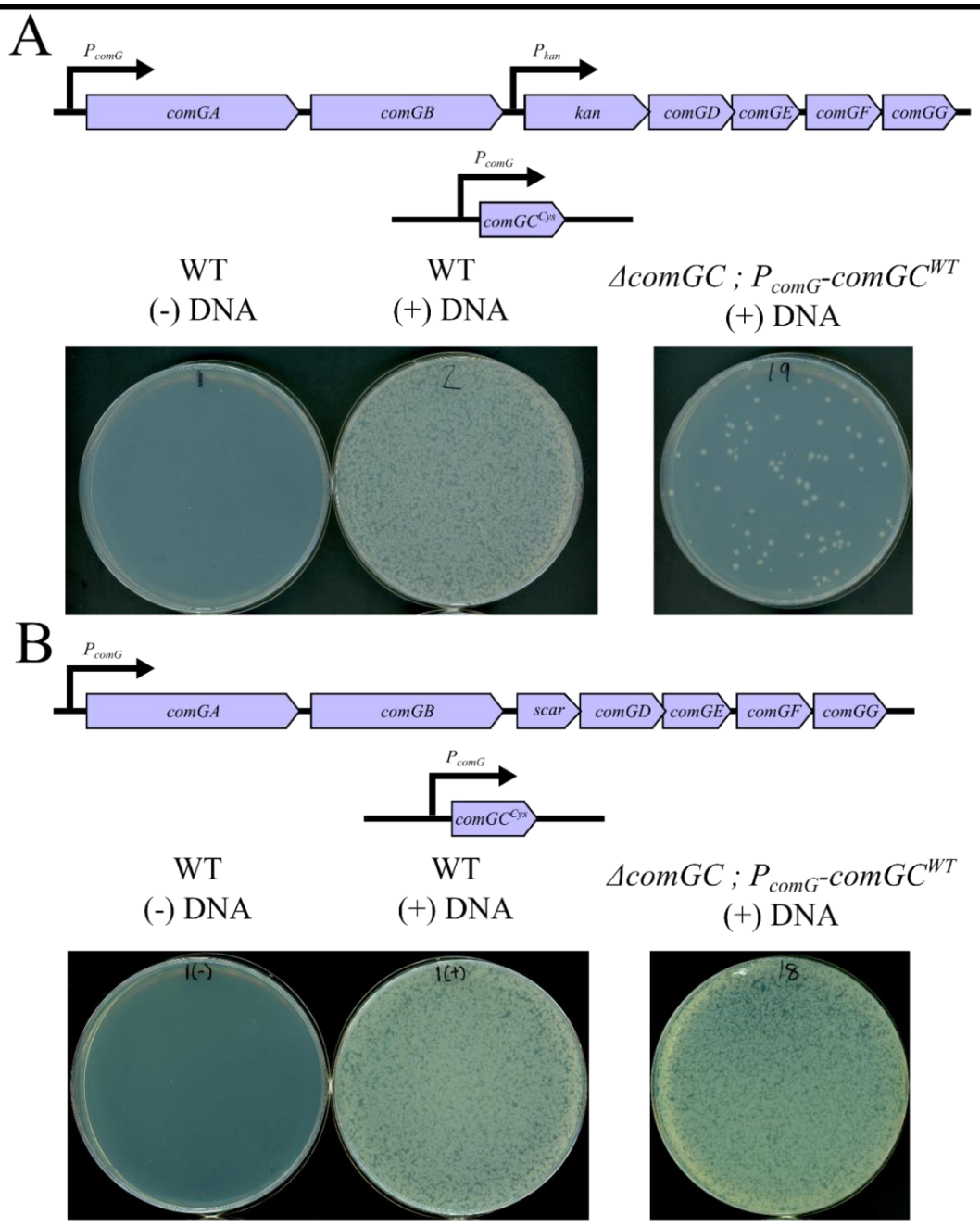
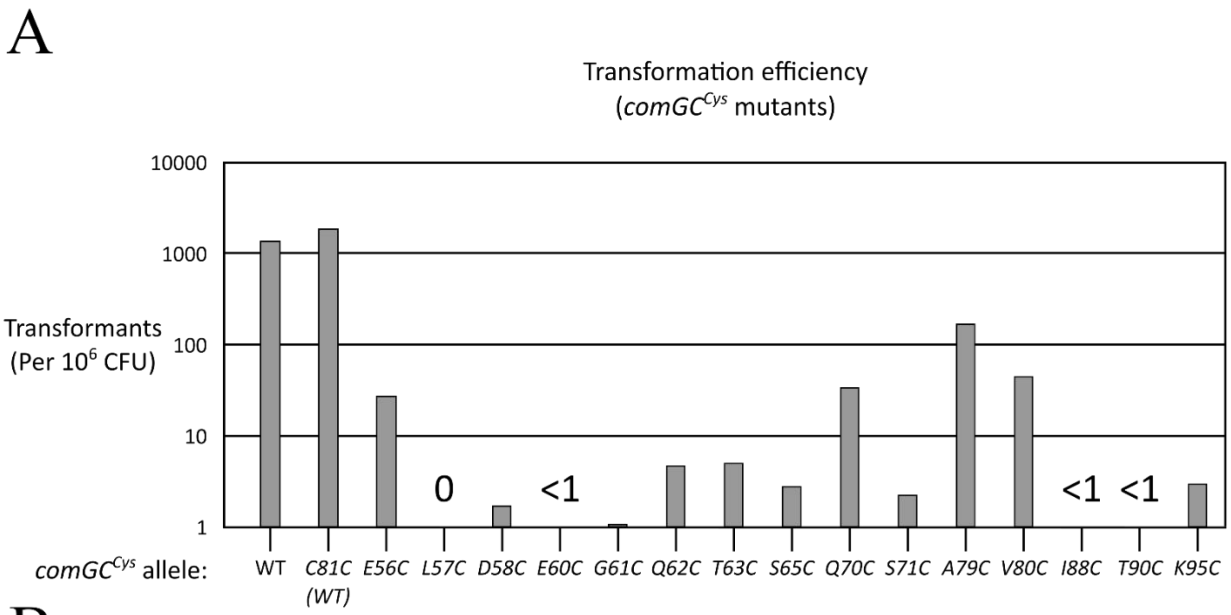


Fig. 9: Cre-recombined Δ*comGC* could be fully complemented by ectopic *comGC*^{WT} expression

(A) Top panel: A schematic of the relevant loci found in the original Δ*comGC* ectopic complementation mutants. The straight black lines represent genomic segments, the purple arrows denote *comG* operon ORFs (or the ectopic complementation *comGC*^{Cys} ORF), and the black arrows show which promoters are driving gene expression at each locus; Bottom panel: Scans of selective LB-Miller agar plates after natural transformation of the indicated strains and plating undiluted cultures.

(B) Top panel: A schematic of the relevant loci found in the Cre-recombinase-recombined Δ*comGC* ectopic complementation mutants. Components are depicted as described in (A); Bottom panel: Same as described in (A).

Figure 10



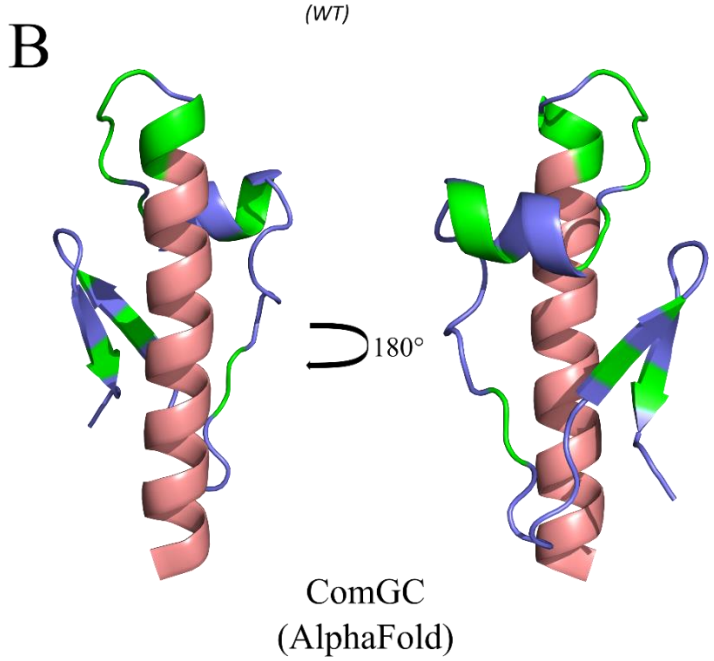


Fig. 10: Ectopic expression of *comGC*^{Cys} alleles could complement for Δ*comGC*

(A) Bar chart displaying the transformation efficiency of the 16 *comGC*^{Cys} alleles capable of complementing for Δ*comGC* in the initial screen, as well as PY79 (WT) and ectopically expressed comGC^{WT} (C81C (WT)). This chart is representative of a single transformation efficiency experiment.

(B) Two views of a cartoon of the highest ranked AlphaFold2 structure of monomeric *Bacillus subtilis* ComGC C-terminal domain, identical to the ComGC C-terminal domain from Fig. 1A. Shown in green are the 9 residues which support cysteine substitution and result in transformation efficiency within 500-fold of wildtype.

Figure 11

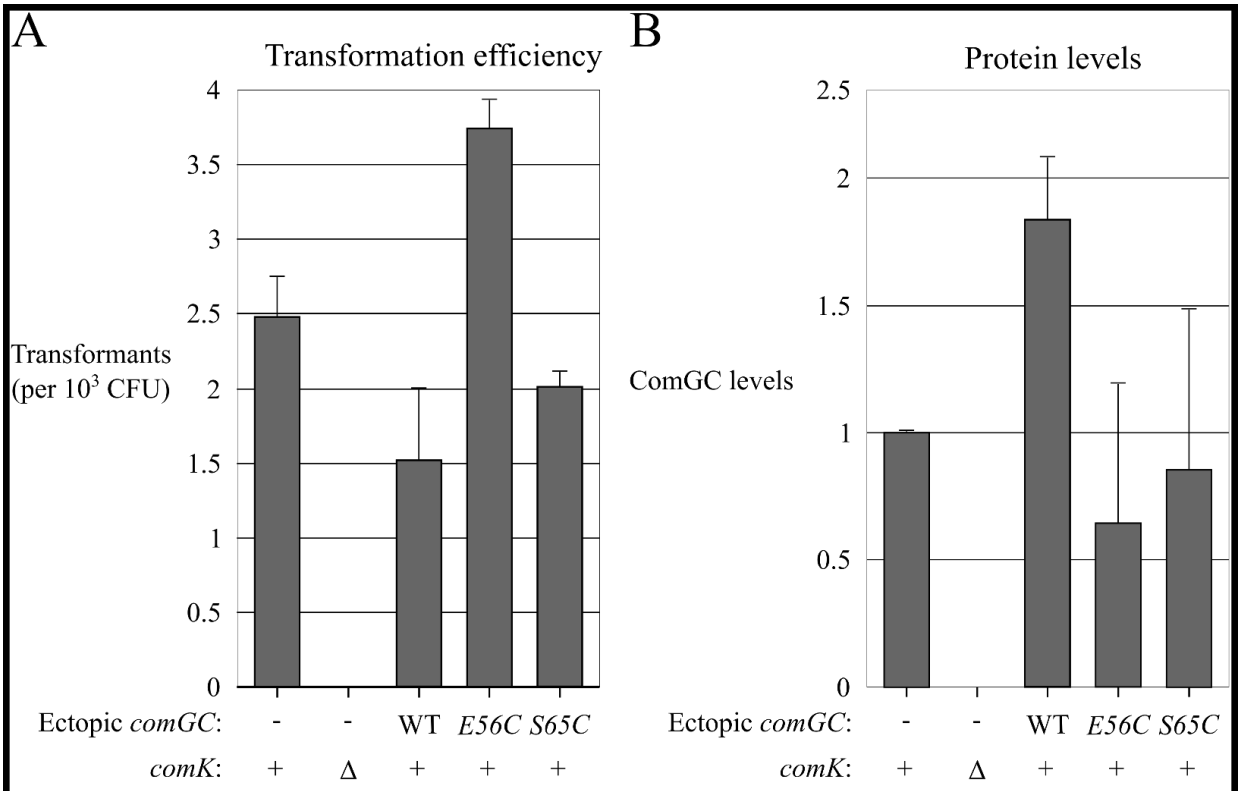


Fig. 11: *B. subtilis* strains ectopically expressing *comGC*^{Cys} are transformable.

(A) Transformation efficiency of *B. subtilis* strains expressing *comGC*^{Cys} from an ectopic locus. All strains carry inducible P_{xyl}-*rbs*-*comK*. The presence and identity of the second *comGC* copy is indicated. The Δ*comK* strain did not have inducer added. The ratio of spectinomycin resistant transformants to total colony forming units was determined. The bar chart represents three biological replicates (error bars represent population standard deviation).

(B) ComGC protein levels from whole cell lysates of each strain referenced in Fig. 11A. Cells from each strain were harvested after one hour of competence induction, total protein for loading normalization was visualized using Ponceau staining, and ComGC levels assessed through Western blotting using anti-ComGC 1° antisera. Normalized ComGC levels were calculated via ImageJ, using total protein as a loading control. Relative ComGC levels are shown relative to P_{xyl}-*rbs*-*comK* and representative of samples taken from three biological replicates.

Figure 12

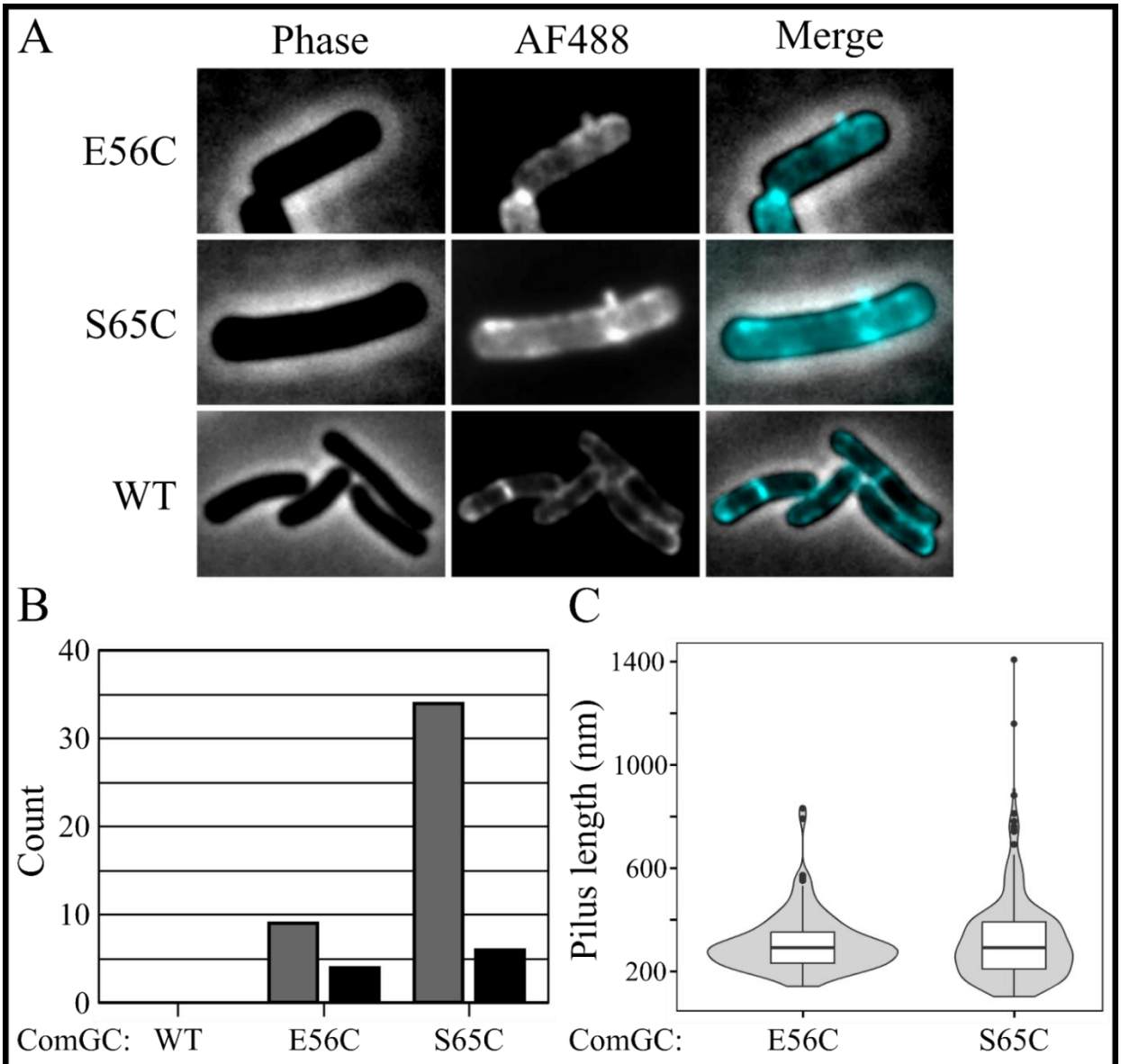


Fig. 12: Strains expressing *comGC*^{Cys} variants produce extracellular filaments composed of ComGC^{Cys} monomers.

(A) Representative images of *B. subtilis* strains producing ComGC^{Cys} extracellular filaments. Competence was induced for one hour, followed by labeling with 25 µg/mL Alexa Fluor 488-maleimide for 20 minutes. Labeled cells were deposited on agar pads and imaged.

(B) Quantification of ComGC pilus production in the same strains and culturing conditions referenced in Fig. 12A. The number of pili (black bars) or pilus-like elements (grey bars) was determined. External foci were defined as any discrete region of fluorescence above background that occurred within 0.5 µm of a cell, whereas pili were defined as any filamentous region of fluorescence contiguous to the boundary of a cell. Quantification was performed on 300 individual cells from each culture.

(C) Quantification of ComGC^{E56C} and ComGC^{S65C} pilus lengths. Violin plots of 133 ComGC^{E56C} pili and 233 ComGC^{S65C} pili measured from cell periphery to tip of pilus. Mean lengths are 310 and 320 nm respectively.

Figure 13

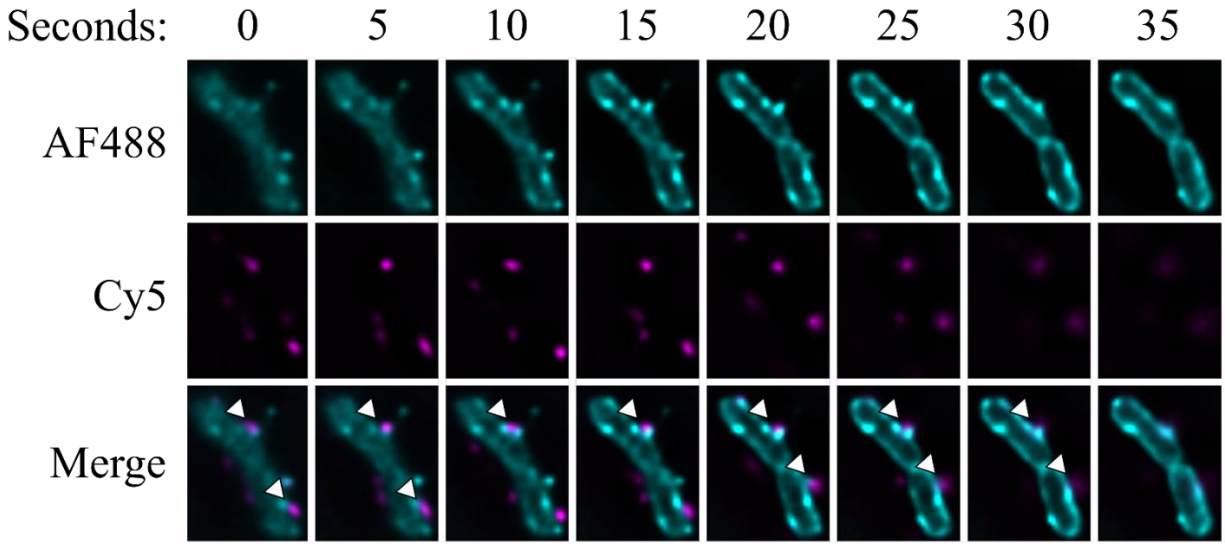


Fig. 13: ComGC pili bind to extracellular DNA.

Representative time series of ComGC pili binding to fluorescent DNA. ComGC^{S65C} pili were labeled as described previously, and fluorescently labeled PCR product was added to the cells prior to deposition on agar pads and imaging. Arrows indicate co-localizing pili and DNA.

Figure 14

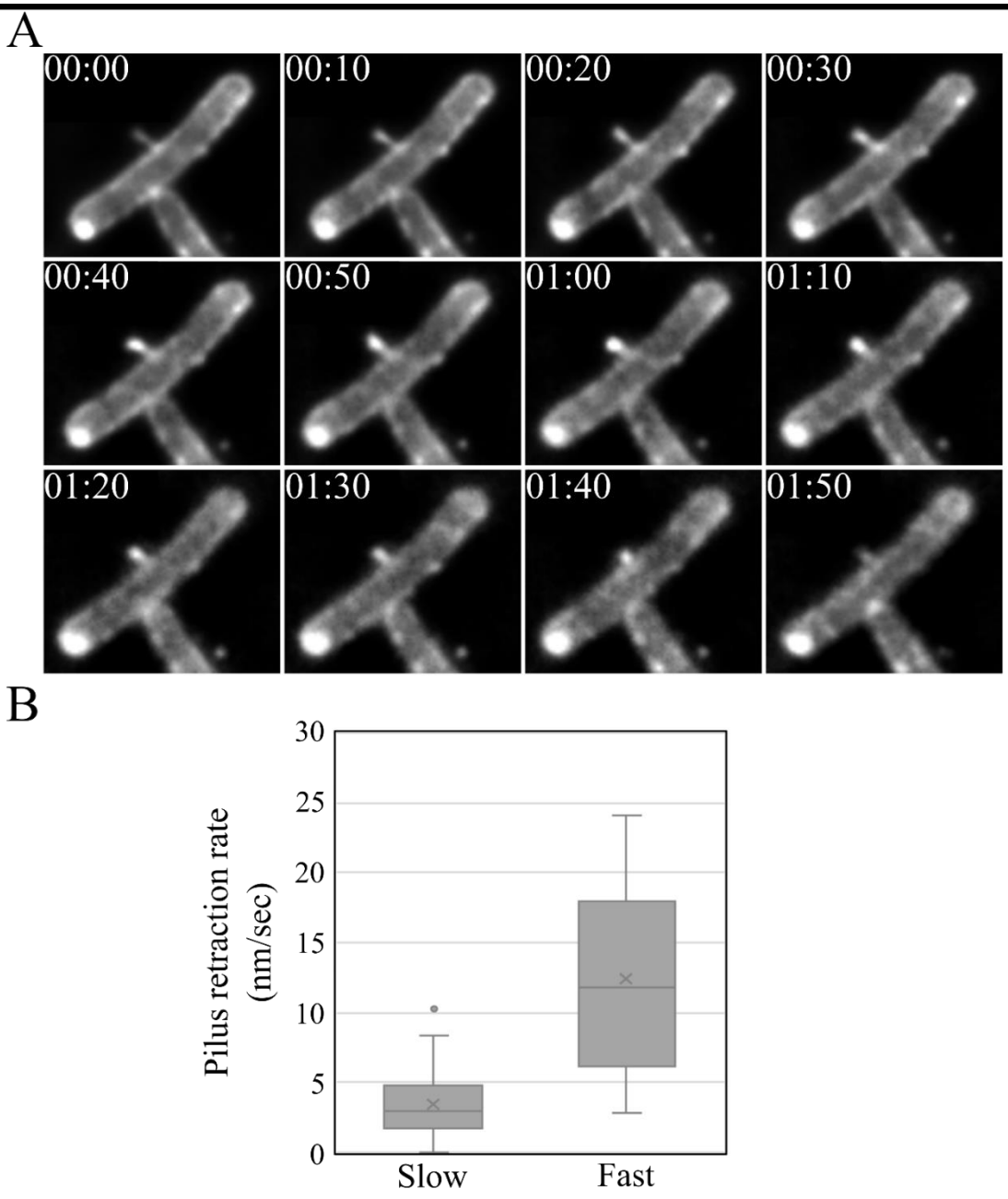


Fig. 14: ComGC pili are retractile.

(A) Representative time series of ComGC pilus retraction. *comGC*^{S65C} was induced, labeled with Alexa Fluor 488-maleimide, and imaged on an agar pad as described previously. Cells were imaged every 5 seconds.

(B) Box-and-whisker plots of ComGC^{S65C} pilus retraction rates drawn from the "slow" population (n = 4 pili) and the "fast" population (n = 6 pili). Individual retraction intervals from data like the example shown in (A) were combined and plotted using Microsoft Excel (n = 22 intervals from each population).

Figure 15

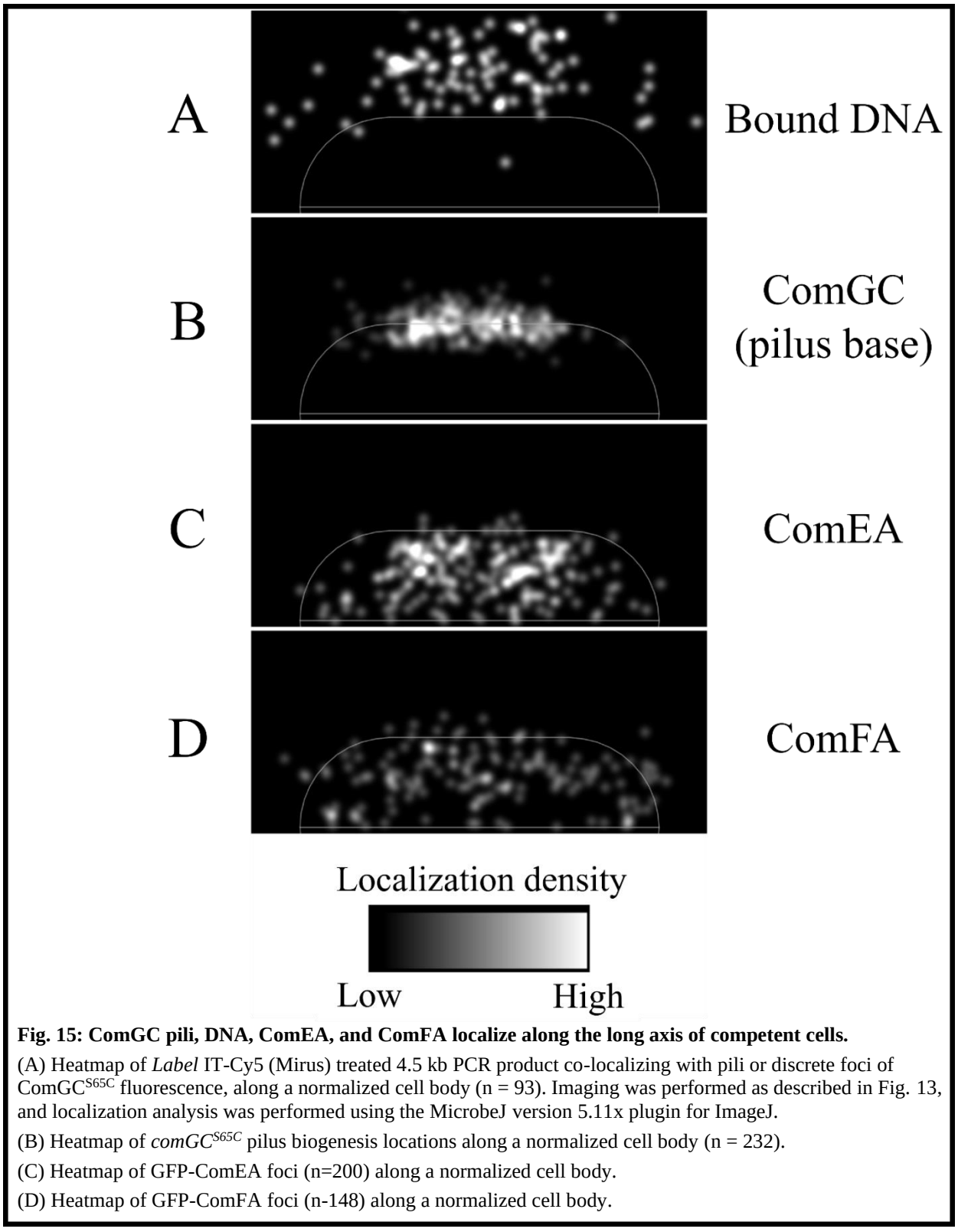


Fig. 15: ComGC pili, DNA, ComEA, and ComFA localize along the long axis of competent cells.

(A) Heatmap of *Label* IT-Cy5 (Mirus) treated 4.5 kb PCR product co-localizing with pili or discrete foci of ComGC^{S65C} fluorescence, along a normalized cell body (n = 93). Imaging was performed as described in Fig. 13, and localization analysis was performed using the MicrobeJ version 5.11x plugin for ImageJ.

(B) Heatmap of *comGC*^{S65C} pilus biogenesis locations along a normalized cell body (n = 232).

(C) Heatmap of GFP-ComEA foci (n=200) along a normalized cell body.

(D) Heatmap of GFP-ComFA foci (n-148) along a normalized cell body.

Tables

Table 1

comGC allele in strain	Transformable?[a]	Transformation efficiency (TFU/CFU)
comGC-C81C (WT)	+	1.9E-03
comGC-E56C	+	2.75E-05
comGC-L57C	+	0.00E+00
comGC-D58C	+	1.72E-06
comGC-H59C	-	N/A
comGC-E60C	+	3.85E-08
comGC-G61C	+	1.06E-06
comGC-Q62C	+	4.73E-06
comGC-T63C	+	4.98E-06
comGC-P64C	-	N/A
comGC-S65C	+	2.77E-06
comGC-L66C	-	N/A
comGC-A67C	-	N/A
comGC-D68C	-	N/A
comGC-L69C	-	N/A
comGC-Q70C	+	3.35E-05
comGC-S71C	+	2.23E-06
comGC-E72C	-	N/A
comGC-G73C	-	N/A
comGC-Y74C	-	N/A
comGC-V75C	-	N/A
comGC-K76C	-	N/A
comGC-K77C	-	N/A
comGC-D78C	-	N/A
comGC-A79C	+	1.69E-04
comGC-V80C	+	4.43E-05
comGC-P82C	-	N/A
comGC-N83C	-	N/A
comGC-G84C	-	N/A
comGC-K85C	-	N/A
comGC-R86C	-	N/A
comGC-I87C	-	N/A
comGC-I88C	+	9.76E-08
comGC-I89C	-	N/A
comGC-T90C	+	6.00E-08
comGC-G91C	-	N/A
comGC-G92C	-	N/A
comGC-E93C	-	N/A

comGC-V94C	-	N/A
comGC-K95C	+	2.95E-06
comGC-V96C	-	N/A
comGC-E97C	-	N/A
comGC-H98C	-	N/A

Transformability of all *comGC^{Cys}* mutants in a Δ*comGC* background

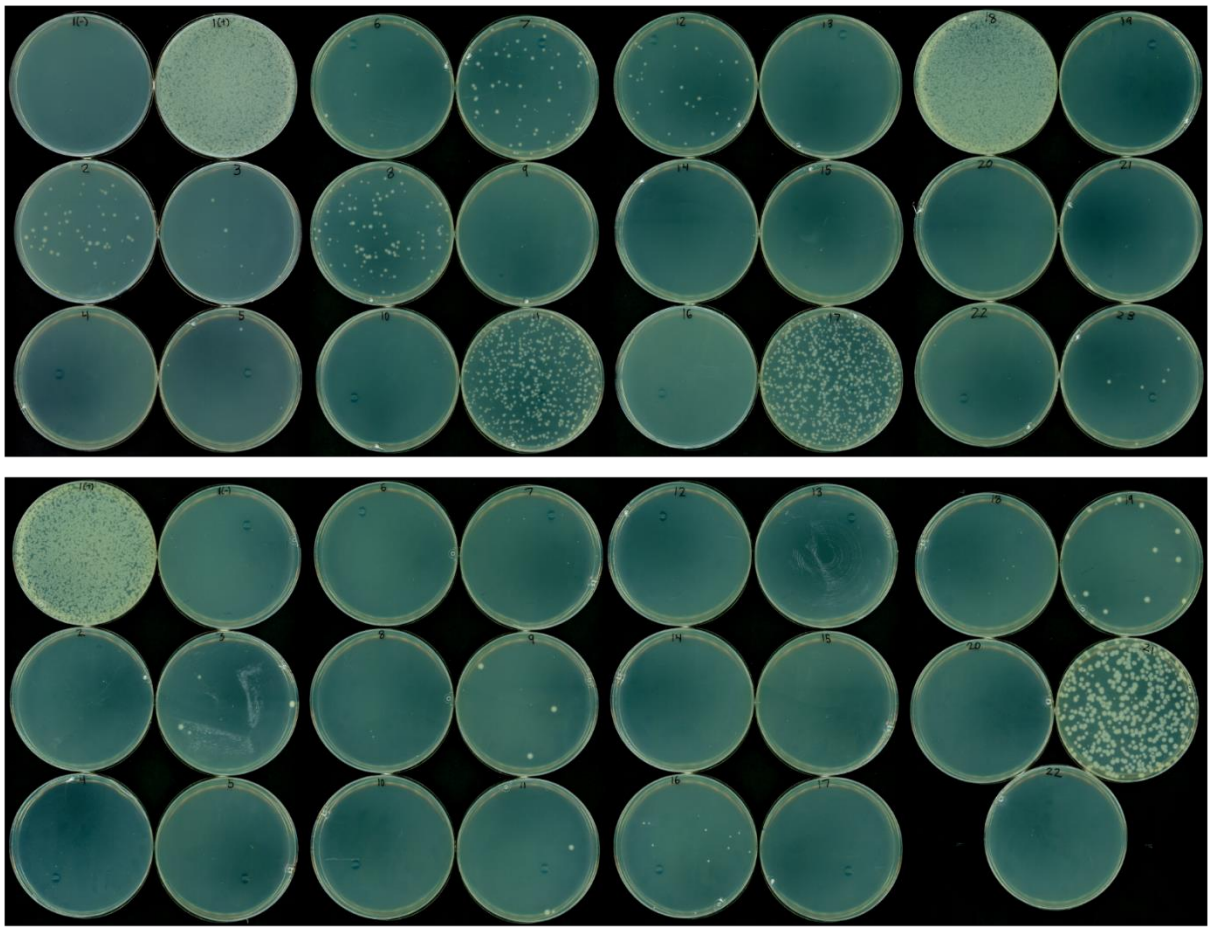
Demonstration of the overall ability for a given *comGC^{Cys}* mutant to support natural transformation in a Δ*comGC* background, as well as the quantitative transformation efficiency of a given mutant that was capable of natural transformation.

[a] A (+) designation shows that a given mutant could produce at least one colony after natural transformation and plating on a selective LB-Miller plate ; a (-) designation shows that a mutant had transformability below the limit of detection in our assay (~ 5 x 10^{-8} transformants/CFU).

Supplementary material

Supplementary figures

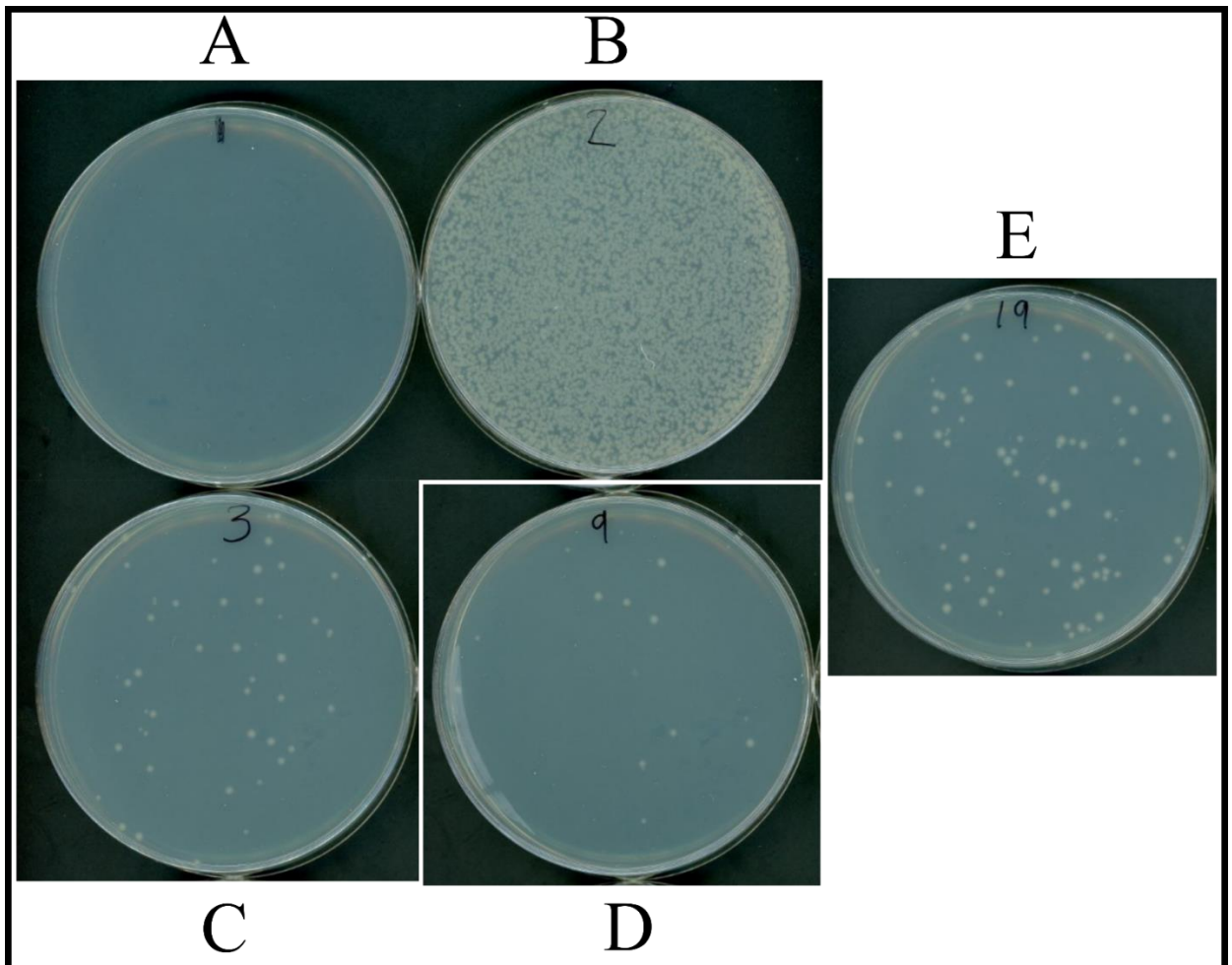
Supplementary Figure 1



Sup. Fig. 1: Original transformation plates of each *comGC*^{Cys} mutant in a Cre-recombinase-recombined Δ*comGC* background

Selective LB-Miller agar plate scans after natural transformation of Cre-recombinase-recombined Δ*comGC* complementation strains after plating undiluted cultures. These data went into the Table 1 column "Transformable?[a]".

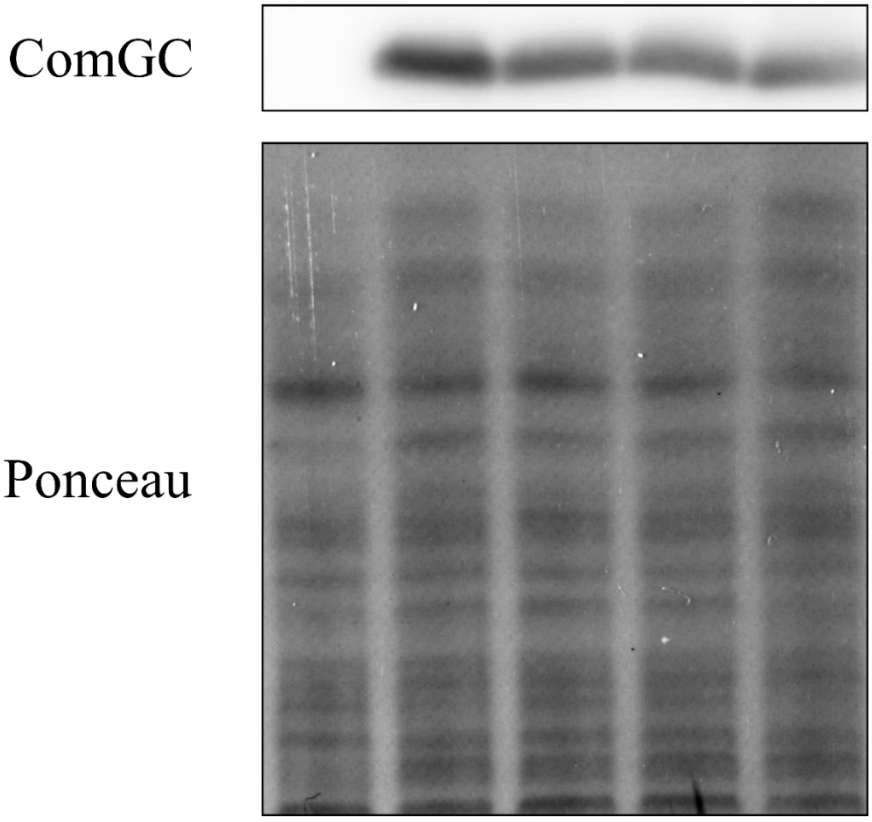
Supplementary Figure 2



Sup. Fig. 2: *comGC*^{Cys} complements poorly for Δ*comGC* on its own

Select ComGC^{Cys} variants support low -efficiency transformation as the sole ComGC variant produced. Transformants resulting from transformation of *B. subtilis* genomic DNA bearing a spectinomycin resistance gene and plating on LB + Spc, using strains (A) PY79 (no DNA) (B) PY79 (C) *lacA::P*$_{comG}$*-comGC*^{E56C} ; Δ*comGC*^{WT} (D) *lacA::P*$_{comG}$*-comGC*^{S65C} ; Δ*comGC*^{WT} (E) *lacA::P*$_{comG}$*-comGC*^{WT} ; Δ*comGC*^{WT}.

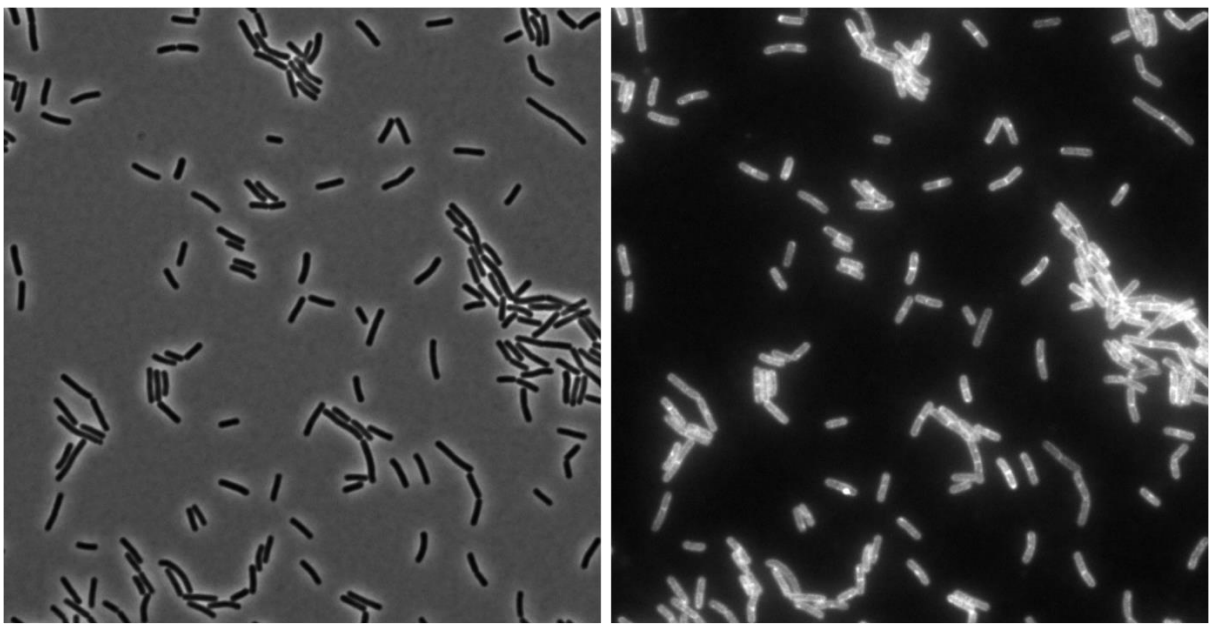
Supplementary Figure 3



Sup. Fig. 3: Example Western blot of ComGC

Representative whole cell lysate ComGC Western blots and Ponceau total protein staining of the following strains cultured until maximum competence was induced, from left to right: *ΔcomK* (uninduced), *comGC*^{WT}, *comGC*^{E56C}, *comGC*^{S65C}, *ΔcomK*. All strains ectopically maintained P_{xyl}-*rbs-comK* and were induced with 0.5% (w/v) xylose unless otherwise noted.

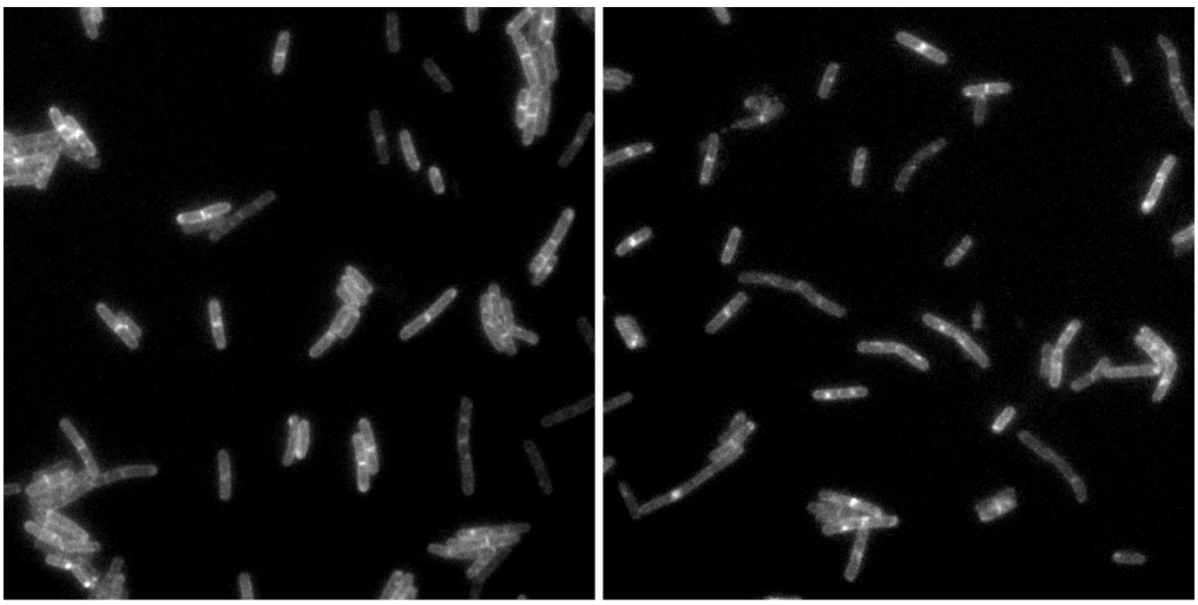
Supplementary Figure 4



Sup. Fig. 4: *comGC*^{WT} does not form pili

Representative imaging field of maleimide-labeled *comGC*^{WT}. No external protruding structures were observed.

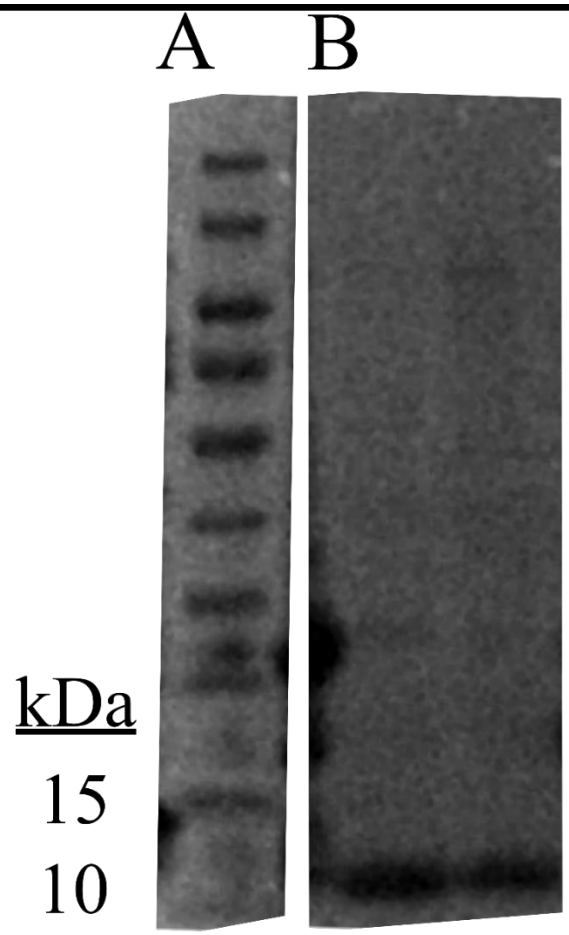
Supplementary Figure 5



Sup. Fig. 5: *comGC*^{Cys} does not form pili without *comGC*^{WT}

Representative imaging field of maleimide-labeled cells with either *comGC*^{E56C} or *comGC*^{S65C} as the sole copy of *comGC* expressed. No external protruding structures were observed in the AF488 channel.

Supplementary Figure 6



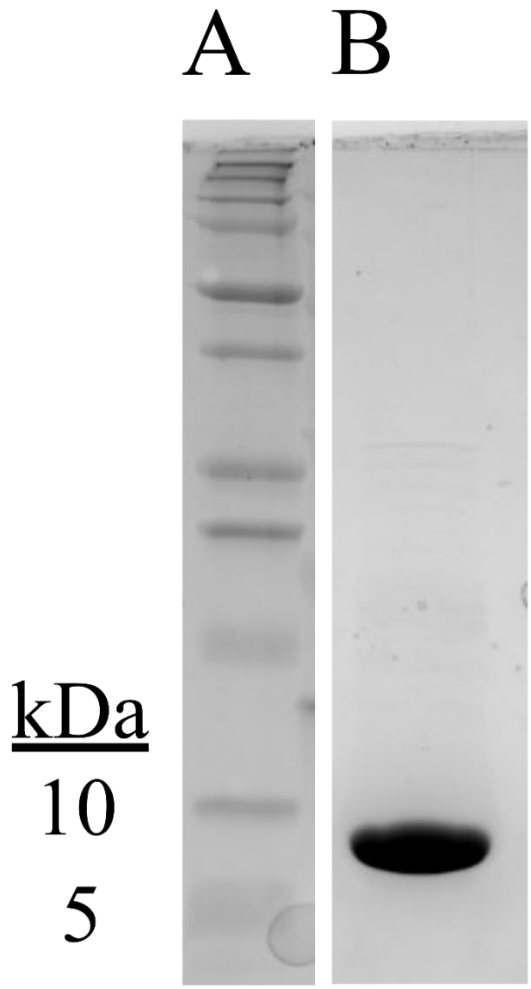
Sup. Fig. 6: ComGC$^{\Delta 1\text{-}30}$ can be efficiently produced

Coomassie stained SDS-PAGE gel after electrophoresis of the ComGC$^{\Delta 1\text{-}30}$ sample following SUMO cleavage and subtractive IMAC

(A) Molecular weight markers, with the 10 and 15 kDa bands labeled

(B) Both ComGC$^{\Delta 1\text{-}30}$ samples

Supplementary Figure 7



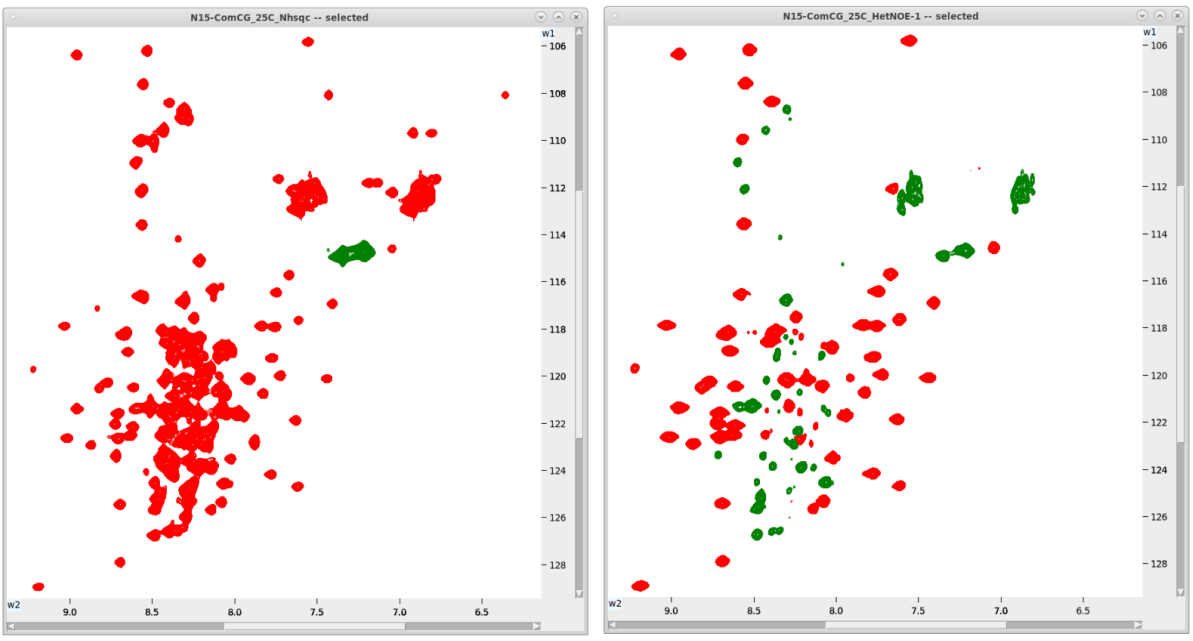
Sup. Fig. 7: ComGC$^{\Delta 1-30}$ can be efficiently purified and highly concentrated

Coomassie stained SDS-PAGE gel after electrophoresis of the ComGC$^{\Delta 1-30}$ sample following all purification steps and concentrations

(A) Molecular weight markers, with the 5 and 10 kDa bands labeled

(B) Combined ComGC$^{\Delta 1-30}$ sample

Supplementary Figure 8



Sup. Fig. 8: NMR spectra of purified ComGC$^{\Delta 1\text{-}30}$

^{15}N HSQC and ^{1}H-^{15}N heteronuclear NOE NMR spectra obtained from ComGC$^{\Delta 1\text{-}30}$. Red and green chemical shifts are positive and negative values, respectively.

(A) ^{15}N HSQC spectrum, note that chemical shifts are highly concentrated

(B) ^{1}H-^{15}N heteronuclear NOE spectrum, note the many negative chemical shifts present

Supplementary tables

Supplementary Table 1

Strains used in Chapter 2

Strain	Genotype	Parent strain	Source, Reference
B. subtilis strain list			
PY79	Prototrophic domesticated laboratory strain	N/A	(179)
bJZ096	*lacA::P_{comG}-comGC^{E56C} (erm)*	PY79	This study
bJZ097	*lacA::P_{comG}-comGC^{L57C} (erm)*	PY79	This study
bJZ098	*lacA::P_{comG}-comGC^{D58C} (erm)*	PY79	This study
bJZ099	*lacA::P_{comG}-comGC^{H59C} (erm)*	PY79	This study
bJZ100	*lacA::P_{comG}-comGC^{G61C} (erm)*	PY79	This study
bJZ101	*lacA::P_{comG}-comGC^{Q62C} (erm)*	PY79	This study
bJZ102	*lacA::P_{comG}-comGC^{T63C} (erm)*	PY79	This study
bJZ103	*lacA::P_{comG}-comGC^{S65C} (erm)*	PY79	This study
bJZ104	*lacA::P_{comG}-comGC^{L66C} (erm)*	PY79	This study
bJZ105	*lacA::P_{comG}-comGC^{A67C} (erm)*	PY79	This study
bJZ106	*lacA::P_{comG}-comGC^{D68C} (erm)*	PY79	This study
bJZ107	*lacA::P_{comG}-comGC^{L69C} (erm)*	PY79	This study
bJZ108	*lacA::P_{comG}-comGC^{Q70C} (erm)*	PY79	This study
bJZ109	*lacA::P_{comG}-comGC^{S71C} (erm)*	PY79	This study
bJZ110	*lacA::P_{comG}-comGC^{E72C} (erm)*	PY79	This study
bJZ111	*lacA::P_{comG}-comGC^{G73C} (erm)*	PY79	This study
bJZ112	*lacA::P_{comG}-comGC^{Y74C} (erm)*	PY79	This study
bJZ113	*lacA::P_{comG}-comGC^{K76C} (erm)*	PY79	This study
bJZ114	*lacA::P_{comG}-comGC^{K77C} (erm)*	PY79	This study
bJZ115	*lacA::P_{comG}-comGC^{D78C} (erm)*	PY79	This study
bJZ116	*lacA::P_{comG}-comGC^{V80C} (erm)*	PY79	This study
bJZ117	*lacA::P_{comG}-comGC^{C81C} (erm)*	PY79	This study
bJZ118	*lacA::P_{comG}-comGC^{N83C} (erm)*	PY79	This study
bJZ119	*lacA::P_{comG}-comGC^{G84C} (erm)*	PY79	This study
bJZ120	*lacA::P_{comG}-comGC^{K85C} (erm)*	PY79	This study
bJZ121	*lacA::P_{comG}-comGC^{R86C} (erm)*	PY79	This study
bJZ122	*lacA::P_{comG}-comGC^{I88C} (erm)*	PY79	This study
bJZ123	*lacA::P_{comG}-comGC^{I89C} (erm)*	PY79	This study
bJZ124	*lacA::P_{comG}-comGC^{G92C} (erm)*	PY79	This study
bJZ125	*lacA::P_{comG}-comGC^{E93C} (erm)*	PY79	This study
bJZ126	*lacA::P_{comG}-comGC^{V94C} (erm)*	PY79	This study
bJZ127	*lacA::P_{comG}-comGC^{K95C} (erm)*	PY79	This study
bJZ128	*lacA::P_{comG}-comGC^{V96C} (erm)*	PY79	This study
bJZ129	*lacA::P_{comG}-comGC^{E97C} (erm)*	PY79	This study
bJZ130	*lacA::P_{comG}-comGC^{H98C} (erm)*	PY79	This study
bJZ133	*lacA::P_{comG}-comGC^{E60C} (erm)*	PY79	This study
bJZ134	*lacA::P_{comG}-comGC^{P64C} (erm)*	PY79	This study
bJZ135	*lacA::P_{comG}-comGC^{V75C} (erm)*	PY79	This study
bJZ136	*lacA::P_{comG}-comGC^{P82C} (erm)*	PY79	This study
bJZ137	*lacA::P_{comG}-comGC^{I87C} (erm)*	PY79	This study
bJZ138	*lacA::P_{comG}-comGC^{T90C} (erm)*	PY79	This study
bJZ139	*lacA::P_{comG}-comGC^{G91C} (erm)*	PY79	This study

bJZ140	*lacA::P_{comG}-comGC^{E56C} (erm) ; comGC::kan*	bJZ096	This study
bJZ141	*lacA::P_{comG}-comGC^{D58C} (erm) ; comGC::kan*	bJZ098	This study
bJZ142	*lacA::P_{comG}-comGC^{H59C} (erm) ; comGC::kan*	bJZ099	This study
bJZ143	*lacA::P_{comG}-comGC^{G61C} (erm) ; comGC::kan*	bJZ100	This study
bJZ144	*lacA::P_{comG}-comGC^{Q62C} (erm) ; comGC::kan*	bJZ101	This study
bJZ145	*lacA::P_{comG}-comGC^{T63C} (erm) ; comGC::kan*	bJZ102	This study
bJZ146	*lacA::P_{comG}-comGC^{S65C} (erm) ; comGC::kan*	bJZ103	This study
bJZ147	*lacA::P_{comG}-comGC^{A67C} (erm) ; comGC::kan*	bJZ105	This study
bJZ148	*lacA::P_{comG}-comGC^{L69C} (erm) ; comGC::kan*	bJZ107	This study
bJZ149	*lacA::P_{comG}-comGC^{Q70C} (erm) ; comGC::kan*	bJZ108	This study
bJZ150	*lacA::P_{comG}-comGC^{S71C} (erm) ; comGC::kan*	bJZ109	This study
bJZ151	*lacA::P_{comG}-comGC^{E72C} (erm) ; comGC::kan*	bJZ110	This study
bJZ152	*lacA::P_{comG}-comGC^{G73C} (erm) ; comGC::kan*	bJZ111	This study
bJZ153	*lacA::P_{comG}-comGC^{K76C} (erm) ; comGC::kan*	bJZ113	This study
bJZ154	*lacA::P_{comG}-comGC^{K77C} (erm) ; comGC::kan*	bJZ114	This study
bJZ155	*lacA::P_{comG}-comGC^{V80C} (erm) ; comGC::kan*	bJZ116	This study
bJZ156	*lacA::P_{comG}-comGC^{C81C} (erm) ; comGC::kan*	bJZ117	This study
bJZ157	*lacA::P_{comG}-comGC^{N83C} (erm) ; comGC::kan*	bJZ118	This study
bJZ158	*lacA::P_{comG}-comGC^{G84C} (erm) ; comGC::kan*	bJZ119	This study
bJZ159	*lacA::P_{comG}-comGC^{K85C} (erm) ; comGC::kan*	bJZ120	This study
bJZ160	*lacA::P_{comG}-comGC^{R86C} (erm) ; comGC::kan*	bJZ121	This study
bJZ161	*lacA::P_{comG}-comGC^{I88C} (erm) ; comGC::kan*	bJZ122	This study
bJZ162	*lacA::P_{comG}-comGC^{I89C} (erm) ; comGC::kan*	bJZ123	This study
bJZ163	*lacA::P_{comG}-comGC^{K95C} (erm) ; comGC::kan*	bJZ127	This study
bJZ164	*lacA::P_{comG}-comGC^{E97C} (erm) ; comGC::kan*	bJZ129	This study
bJZ165	*lacA::P_{comG}-comGC^{H98C} (erm) ; comGC::kan*	bJZ130	This study
bJZ166	*lacA::P_{comG}-comGC^{P64C} (erm) ; comGC::kan*	bJZ134	This study
bJZ167	*lacA::P_{comG}-comGC^{P82C} (erm) ; comGC::kan*	bJZ136	This study
bJZ168	*lacA::P_{comG}-comGC^{I87C} (erm) ; comGC::kan*	bJZ137	This study
bJZ169	*lacA::P_{comG}-comGC^{T90C} (erm) ; comGC::kan*	bJZ138	This study
bJZ170	*lacA::P_{comG}-comGC^{G91C} (erm) ; comGC::kan*	bJZ139	This study
bJZ171	*lacA::P_{comG}-comGC^{A79C} (erm)*	PY79	This study
bJZ172	*lacA::P_{comG}-comGC^{L57C} (erm) ; comGC::kan*	bJZ097	This study
bJZ173	*lacA::P_{comG}-comGC^{D68C} (erm) ; comGC::kan*	bJZ106	This study
bJZ174	*lacA::P_{comG}-comGC^{Y74C} (erm) ; comGC::kan*	bJZ112	This study
bJZ175	*lacA::P_{comG}-comGC^{D78C} (erm) ; comGC::kan*	bJZ115	This study
bJZ176	*lacA::P_{comG}-comGC^{G92C} (erm) ; comGC::kan*	bJZ124	This study
bJZ177	*lacA::P_{comG}-comGC^{E93C} (erm) ; comGC::kan*	bJZ125	This study
bJZ178	*lacA::P_{comG}-comGC^{V94C} (erm) ; comGC::kan*	bJZ126	This study
bJZ179	*lacA::P_{comG}-comGC^{V96C} (erm) ; comGC::kan*	bJZ128	This study
bJZ180	*lacA::P_{comG}-comGC^{E60C} (erm) ; comGC::kan*	bJZ133	This study
bJZ181	*lacA::P_{comG}-comGC^{V75C} (erm) ; comGC::kan*	bJZ135	This study
bJZ182	*lacA::P_{comG}-comGC^{A79C} (erm) ; comGC::kan*	bJZ171	This study
bJZ183	*lacA::P_{comG}-comGC^{L66C} (erm) ; comGC::kan*	bJZ104	This study
bJZ238	*comGC::kan*	PY79	This study
bJZ363	*amyE::$P_{spank\text{-}hy}$-comG (spc)*	PY79	This study
bJZ447	*amyE::$P_{spank\text{-}hy}$ rapI-catR ; ICEBs1 Δ(rapI-phrI)::kan ; ΔattR::tet ; ΔnicK(unmarked) (oriT-) ; pDR244*	ELC453	This study
bJZ450	*lacA::P_{comG}-comGC^{E56C} (erm) ; comGC markerless deletion*	bJZ140	This study
bJZ451	*lacA::P_{comG}-comGC^{D58C} (erm) ; comGC markerless deletion*	bJZ141	This study
bJZ452	*lacA::P_{comG}-comGC^{H59C} (erm) ; comGC markerless deletion*	bJZ142	This study
bJZ453	*lacA::P_{comG}-comGC^{G61C} (erm) ; comGC markerless deletion*	bJZ143	This study
bJZ454	*lacA::P_{comG}-comGC^{Q62C} (erm) ; comGC markerless deletion*	bJZ144	This study
bJZ455	*lacA::P_{comG}-comGC^{T63C} (erm) ; comGC markerless deletion*	bJZ145	This study

bJZ456	*lacA::P_{comG}-comGC^{S65C} (erm) ; comGC markerless deletion*	bJZ146	This study
bJZ457	*lacA::P_{comG}-comGC^{A67C} (erm) ; comGC markerless deletion*	bJZ147	This study
bJZ458	*lacA::P_{comG}-comGC^{L69C} (erm) ; comGC markerless deletion*	bJZ148	This study
bJZ459	*lacA::P_{comG}-comGC^{Q70C} (erm) ; comGC markerless deletion*	bJZ149	This study
bJZ460	*lacA::P_{comG}-comGC^{S71C} (erm) ; comGC markerless deletion*	bJZ150	This study
bJZ461	*lacA::P_{comG}-comGC^{E72C} (erm) ; comGC markerless deletion*	bJZ151	This study
bJZ462	*lacA::P_{comG}-comGC^{G73C} (erm) ; comGC markerless deletion*	bJZ152	This study
bJZ463	*lacA::P_{comG}-comGC^{K76C} (erm) ; comGC markerless deletion*	bJZ153	This study
bJZ464	*lacA::P_{comG}-comGC^{K77C} (erm) ; comGC markerless deletion*	bJZ154	This study
bJZ465	*lacA::P_{comG}-comGC^{V80C} (erm) ; comGC markerless deletion*	bJZ155	This study
bJZ466	*lacA::P_{comG}-comGC^{C81C} (erm) ; comGC markerless deletion*	bJZ156	This study
bJZ467	*lacA::P_{comG}-comGC^{N83C} (erm) ; comGC markerless deletion*	bJZ157	This study
bJZ468	*lacA::P_{comG}-comGC^{G84C} (erm) ; comGC markerless deletion*	bJZ158	This study
bJZ469	*lacA::P_{comG}-comGC^{K85C} (erm) ; comGC markerless deletion*	bJZ159	This study
bJZ470	*lacA::P_{comG}-comGC^{R86C} (erm) ; comGC markerless deletion*	bJZ160	This study
bJZ471	*lacA::P_{comG}-comGC^{I88C} (erm) ; comGC markerless deletion*	bJZ161	This study
bJZ472	*lacA::P_{comG}-comGC^{I89C} (erm) ; comGC markerless deletion*	bJZ162	This study
bJZ473	*lacA::P_{comG}-comGC^{K95C} (erm) ; comGC markerless deletion*	bJZ163	This study
bJZ474	*lacA::P_{comG}-comGC^{E97C} (erm) ; comGC markerless deletion*	bJZ164	This study
bJZ475	*lacA::P_{comG}-comGC^{H98C} (erm) ; comGC markerless deletion*	bJZ165	This study
bJZ476	*lacA::P_{comG}-comGC^{P64C} (erm) ; comGC markerless deletion*	bJZ166	This study
bJZ477	*lacA::P_{comG}-comGC^{P82C} (erm) ; comGC markerless deletion*	bJZ167	This study
bJZ478	*lacA::P_{comG}-comGC^{I87C} (erm) ; comGC markerless deletion*	bJZ168	This study
bJZ479	*lacA::P_{comG}-comGC^{T90C} (erm) ; comGC markerless deletion*	bJZ169	This study
bJZ480	*lacA::P_{comG}-comGC^{G91C} (erm) ; comGC markerless deletion*	bJZ170	This study
bJZ481	*lacA::P_{comG}-comGC^{L57C} (erm) ; comGC markerless deletion*	bJZ172	This study
bJZ482	*lacA::P_{comG}-comGC^{D68C} (erm) ; comGC markerless deletion*	bJZ173	This study
bJZ483	*lacA::P_{comG}-comGC^{Y74C} (erm) ; comGC markerless deletion*	bJZ174	This study
bJZ484	*lacA::P_{comG}-comGC^{D78C} (erm) ; comGC markerless deletion*	bJZ175	This study
bJZ485	*lacA::P_{comG}-comGC^{G92C} (erm) ; comGC markerless deletion*	bJZ176	This study
bJZ486	*lacA::P_{comG}-comGC^{E93C} (erm) ; comGC markerless deletion*	bJZ177	This study
bJZ487	*lacA::P_{comG}-comGC^{V94C} (erm) ; comGC markerless deletion*	bJZ178	This study
bJZ488	*lacA::P_{comG}-comGC^{V96C} (erm) ; comGC markerless deletion*	bJZ179	This study
bJZ489	*lacA::P_{comG}-comGC^{E60C} (erm) ; comGC markerless deletion*	bJZ180	This study
bJZ490	*lacA::P_{comG}-comGC^{V75C} (erm) ; comGC markerless deletion*	bJZ181	This study
bJZ491	*lacA::P_{comG}-comGC^{A79C} (erm) ; comGC markerless deletion*	bJZ182	This study
bJZ492	*lacA::P_{comG}-comGC^{L66C} (erm) ; comGC markerless deletion*	bJZ183	This study
BKK24710	*comGC::kan*	168	(109)
ELC453	*amyE::$P_{spank\text{-}hy}$ rapI-catR ; ICEBs1 Δ(rapI-phrI)::kan ; ΔattR::tet ; ΔnicK(unmarked) (oriT-)*	JH642	A. Grossman
E. coli strain list			
BOSE1620	pBOSE1620 (*amyE::$P_{spank\text{-}hy}$-comG (spc)*)	DH5-α	B. Bose

Supplementary Table 2

Plasmids used in Chapter 2

Plasmid	Description	Source
pBOSE1620	*amyE::$P_{spank\text{-}hy}$-comG (spc)*	B. Bose
pKRH83	*lacA::P_{comG}-comGC (erm)*	D. Kearns
pDR244	*P_{spac}-cre (spc)*	D. Rudner

Supplementary Table 3

DNA oligos used in Chapter 2

Oligo	Sequence (5' to 3')	Source
U-F	CCAGCAACGCGGCCTTTTTACG	Sigma-Aldrich
U-R	GCTGCGCCTTATCCGGTAACTATCG	Sigma-Aldrich
C-F-56	atgactgcatttTGTcttgatcatgaaggac	Eurofins Genomics
C-R-56	tcatgatcaagACAaaatgcagtcatttgtgc	Eurofins Genomics
C-F-57	ctgcatttgagTGTgatcatgaaggacaaac	Eurofins Genomics
C-R-57	cttcatgatcACActcaaatgcagtcatttgtg	Eurofins Genomics
C-F-58	gagcttTGTcatgaaggacaaactccgag	Eurofins Genomics
C-R-58	tttgtccttcatgACAaagctcaaatgcagtc	Eurofins Genomics
C-F-59	tgagcttgatTGTgaaggacaaactcc	Eurofins Genomics
C-R-59	ttgtccttcACAatcaagctcaaatgcag	Eurofins Genomics
C-F-60	gcttgatcatTGTggacaaactccgag	Eurofins Genomics
C-R-60	gtttgtccACAatgatcaagctcaaatgc	Eurofins Genomics
C-F-61	tcatgaaTGTcaaactccgagccttg	Eurofins Genomics
C-R-61	ctcggagtttgACAttcatgatcaagctc	Eurofins Genomics
C-F-62	atgaaggaTGTactccgagccttgc	Eurofins Genomics
C-R-62	gctcggagtACAtccttcatgatcaag	Eurofins Genomics
C-F-63	gacaaTGTccgagccttgccgatttac	Eurofins Genomics
C-R-63	caaggctcggACAttgtccttcatgatc	Eurofins Genomics
C-F-64	ggacaaactTGTagccttgccgatttac	Eurofins Genomics
C-R-64	ggcaaggctACAagtttgtccttcatg	Eurofins Genomics
C-F-65	aactccgTGCcttgccgatttacag	Eurofins Genomics
C-R-65	gcaagGCAcggagtttgtccttc	Eurofins Genomics
C-F-66	ctccgagcTGTgccgatttacagtc	Eurofins Genomics
C-R-66	taaatcggcACAgctcggagtttgtcc	Eurofins Genomics
C-F-67	cgagccttTGCgatttacagtcagag	Eurofins Genomics
C-R-67	ctgtaaatcGCAaaggctcggagtttg	Eurofins Genomics
C-F-68	gccttgccTGTttacagtcagagggctatg	Eurofins Genomics
C-R-68	actgtaaACAggcaaggctcggagtttg	Eurofins Genomics
C-F-69	ccttgccgatTGTcagtcagagggctatg	Eurofins Genomics
C-R-69	actgACAatcggcaaggctcggag	Eurofins Genomics
C-F-70	gccgatttaTGTtcagagggctatgtg	Eurofins Genomics
C-R-70	ccctctgaACAtaaatcggcaaggctc	Eurofins Genomics
C-F-71	ccgatttacagTGTgagggctatgtgaaaaag	Eurofins Genomics
C-R-71	tagccctcACActgtaaatcggcaag	Eurofins Genomics
C-F-72	acagtcaTGTggctatgtgaaaaaggatgc	Eurofins Genomics
C-R-72	tcacatagccACAtgactgtaaatcggcaag	Eurofins Genomics
C-F-73	acagtcagagTGCtatgtgaaaaaggatgc	Eurofins Genomics
C-R-73	ccttttttcacataGCActctgactgtaaatcg	Eurofins Genomics
C-F-74	cagagggcTGTgtgaaaaaggatgc	Eurofins Genomics
C-R-74	ccttttttcacACAgccctctgactgtaaatc	Eurofins Genomics
C-F-75	gtcagagggctatTGTaaaaaggatgctgtc	Eurofins Genomics
C-R-75	gcatccttttttACAatagccctctgactg	Eurofins Genomics
C-F-76	gggctatgtgTGTaaggatgctgtctg	Eurofins Genomics
C-R-76	gcatccttACAcacatagccctctgac	Eurofins Genomics
C-F-77	gctatgtgaaaTGTgatgctgtctgtcc	Eurofins Genomics
C-R-77	cagcatcACAtttcacatagccctctg	Eurofins Genomics
C-F-78	tgtgaaaaagTGTgctgtctgtccaaatg	Eurofins Genomics
C-R-78	cagacagcACActttttcacatagccctc	Eurofins Genomics

C-F-79	aggatTGTgtctgtccaaatggtaagc	Eurofins Genomics
C-R-79	atttggacagacACAatccttttcacatagc	Eurofins Genomics
C-F-80	ggatgctTGCtgtccaaatggtaagc	Eurofins Genomics
C-R-80	catttggacaGCAagcatccttttcac	Eurofins Genomics
C-F-81	gctgtcTGTccaaatggtaagcgcattatc	Eurofins Genomics
C-R-81	accatttggACAgacagcatccttttc	Eurofins Genomics
C-F-82	gctgtctgtTGTaatggtaagcgcattatc	Eurofins Genomics
C-R-82	gcttaccattACAacagacagcatccttttc	Eurofins Genomics
C-F-83	gtctgtccaTGTggtaagcgcattatc	Eurofins Genomics
C-R-83	cgcttaccACAtggacagacagcatc	Eurofins Genomics
C-F-84	gtctgtccaaatTGTaagcgcattatcatcac	Eurofins Genomics
C-R-84	tgcgcttACAatttggacagacagc	Eurofins Genomics
C-F-85	gtccaaatggtTGTcgcattatcatcacc	Eurofins Genomics
C-R-85	aatgcgACAaccatttggacagacagc	Eurofins Genomics
C-F-86	gTGCattatcatcaccggcggag	Eurofins Genomics
C-R-86	ccggtgatgataatGCActtaccatttggacag	Eurofins Genomics
C-F-87	cgcTGTatcatcaccggcggag	Eurofins Genomics
C-R-87	ccggtgatgatACAgcgcttaccatttgg	Eurofins Genomics
C-F-88	gcattTGCatcaccggcggagaag	Eurofins Genomics
C-R-88	gccggtgatGCAaatgcgcttaccatttg	Eurofins Genomics
C-F-89	gcattatcTGCaccggcggagaagttaag	Eurofins Genomics
C-R-89	gccggtGCAgataatgcgcttacc	Eurofins Genomics
C-F-90	gcattatcatcTGCggcggagaagttaag	Eurofins Genomics
C-R-90	tccgccGCAgatgataatgcgcttac	Eurofins Genomics
C-F-91	atcatcaccTGCggagaagttaaggttg	Eurofins Genomics
C-R-91	tccGCAggtgatgataatgcgcttac	Eurofins Genomics
C-F-92	caccggcTGTgaagttaaggttgaac	Eurofins Genomics
C-R-92	ccttaacttcACAgccggtgatgataatg	Eurofins Genomics
C-F-93	ccggcggaTGTgttaaggttgaacattaaattaaac	Eurofins Genomics
C-R-93	cttaacACAtccgccggtgatgataatg	Eurofins Genomics
C-F-94	ggcggagaaTGTaaggttgaacattaaattaaacg	Eurofins Genomics
C-R-94	tcaaccttACAttctccgccggtgatg	Eurofins Genomics
C-F-95	ggcggagaagttTGTgttgaacattaaattaaacg	Eurofins Genomics
C-R-95	ttcaacACAaacttctccgccggtg	Eurofins Genomics
C-F-96	ggcggagaagttaagTGTgaacattaaattaaacg	Eurofins Genomics
C-R-96	ttcACActtaacttctccgccggtg	Eurofins Genomics
C-F-97	gttTGTcattaaattaaacgaggagaagggctagc	Eurofins Genomics
C-R-97	tctcctcgtttaatttaatgACAaaccttaacttctcc	Eurofins Genomics
C-F-98	gttgaaTGTtaaattaaacgaggagaagggctagc	Eurofins Genomics
C-R-98	cttctcctcgtttaatttaACAttcaaccttaacttctcc	Eurofins Genomics
oJZ149	CGGGAGCTTTTCACATACAGC	Sigma-Aldrich
oJZ150	CAGCACAAACAGCCACAGG	Sigma-Aldrich
oJZ151	TGATCCTGCCTCGCTTGTAC	Sigma-Aldrich
oJZ152	GACGTCAGGTGGCACTTTTC	Sigma-Aldrich
oJZ153	TCCGTTGAGAAAGATACTGG	Sigma-Aldrich
oJZ223	TTCGGGTCGATCATGCCTTG	Sigma-Aldrich
oJZ225	AACCACTGCTGGGAAACAGG	Sigma-Aldrich

Supplementary Table 4

Strains used in Chapter 3

Strain	Genotype	Parent strain	Source, Reference
DH5α	F$^-$ *endA1 glnV44 thi1 recA1 relA1 gyrA96 deoR nupG purB2 0* φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17($r_K^- m_K^+$), λ$^-$		(180)
eJZ040	pJZ026 (*yhdGH::$P_{xyl(-CRE)}$ (kan)*)	DH5α	This work
eJZ064	pJZ049 (*yhdGH::$P_{xyl(-CRE)}$-rbs-comGC (kan)*)	DH5α	This work
eJZ095	pJZ069 (*yhdGH::P_{xyl}-rbs-comK (kan)*)	DH5α	This work
PY79	Prototrophic domesticated laboratory strain		(181)
bBB036	*comK::spc*	PY79	This work
bBB050	*amyE::P_{comE}-gfp-comEA (cat)*	PY79	This work
bSC050	*yvbJ::P_{comF}-gfp-comFA (erm)*	PY79	(182)
bJZ083	*comGC::kan*	168	(182)
bJZ096	*lacA::P_{comG}-comGC^{E56C} (erm)*	PY79	This work
bJZ103	*lacA::P_{comG}-comGC^{S65C} (erm)*	PY79	This work
bJZ117	*lacA::P_{comG}-comGC^{WT} (erm)*	PY79	This work
bJZ140	*lacA::P_{comG}-comGC^{E56C} (erm) ; comGC::kan*	bJZ096	This work
bJZ146	*lacA::P_{comG}-comGC^{S65C} (erm) ; comGC::kan*	bJZ103	This work
bJZ156	*lacA::P_{comG}-comGC^{WT} (erm) ; comGC::kan*	bJZ117	This work
bJZ405	*yhdGH::P_{xyl}-rbs-comK (kan)*	PY79	This work
bJZ407	*yhdGH::P_{xyl}-rbs-comK (kan) ; comK::spc*	bJZ405	This work
bJZ418	*yhdGH::P_{xyl}-rbs-comK (kan) ; amyE::P_{comE}-gfp-comEA (cat)*	bJZ405	This work
bJZ419	*yhdGH::P_{xyl}-rbs-comK (kan) ; lacA::P_{comG}-comGC^{E56C} (erm)*	bJZ405	This work
bJZ423	*yhdGH::P_{xyl}-rbs-comK (kan) ; lacA::P_{comG}-comGC^{S65C} (erm)*	bJZ405	This work
bJZ427	*yhdGH::P_{xyl}-rbs-comK (kan) ; lacA::P_{comG}-comGC^{WT} (erm)*	bJZ405	This work

Supplementary Table 5

Plasmids used in Chapter 3

Plasmid	Description	Source
pBB401	*amyE::PcomE-gfp-comEA(cat)*	This work
pBS0E-XylR-P$_{xylA\ (V2)}$	Empty replicative *B. subtilis* plasmid carrying *xylR-P* $_{xylA\ (V2)}$	(145)
pKL147	*dnaX*$^{3'}$*-gfp*	(183)
pKRH83	*lacA::P*$_{comG}$*-comGC*	D. Kearns
pNS008	*amyE::P*$_{SpoIIQ}$*-yfp (cat)*	D. Rudner
pJZ026	*yhdGH::P*$_{xyl(-CRE)}$ *(kan)*	This work
pJZ049	*yhdGH::P*$_{xyl(-CRE)}$*-rbs-comGC (kan)*	This work
pJZ069	*yhdGH::P*$_{xyl}$*-rbs-comK (kan)*	This work

Supplementary Table 6

DNA oligos used in Chapter 3

Primer	Sequence*
oBMB116	GGAATTCCGCTTACGATGTCTGATGCAACC
oBMB117	CGTAAGCTTCGCTTTCGTCATATTTTATTTTC
oBMB127	GGCAAGCTTACATAAGGAGGAACTACTATGAGTAAAGGAGAAGAAC
oBMB128	CGGCTCGAGTTTGTATAGTTCATCCATGC
oBOSE397	CCAGCCAATATTCACTTGCTGCGG
oBOSE400	CTCCCTGACCGCTCGGCTTCC
oJZ149	GCTGCGCCTTATCCGGTAACTATCG
oJZ150	CCAGCAACGCGGCCTTTTTACG
oJZ151	*ATGACTGCATTTTGTCTTGATCATGA*AGGAC
oJZ152	*TCATGATCAAGACAAAATGCAGTCAT*TTGTGC
oJZ167	*GCATGCATTAGCTAGCATTA*CTCGAGACACTAGTATAAGC
oJZ168	*ATTAAGCGCGGCGGGTGT*G
oJZ169	*TAATGCTAGCTAATGCATGC*TTAGTTTGTTGCCCAAAC
oJZ170	*ACACCCGCCGCGCTTAAT*G
oJZ175	*GCCTTGCCTGTTTACAGT*CAGAGGGCTATG
oJZ176	*ACTGTAAACAGGCAAGGC*TCGGAGTTTG
oJZ199	TGGGCAACAAACTAAGCATGGACAATATGCGTTGAAAGGAGA
oJZ200	CGAGTAATGCTAGCTAATGCATGAGTGATCACCTCATCATTCAATTAT
oJZ201	*GCTGTCTGTCCAAATGGT*AAGCGCATTATC
oJZ202	*ACCATTTGGACAGACAGC*ATCCTTTTTC
oJZ212	TGATCCTGCCTCGCTTGTAC
oJZ213	GACGTCAGGTGGCACTTTTC
oJZ214	TCCGTTGAGAAAGATACTGG
oJZ223	TTCGGGTCGATCATGCCTTG
oJZ225	AACCACTGCTGGGAAACAGG
oJZ257	AGGAGGCTAGCCTATGAATGAGAAAGGATTTACAC
oJZ258	*GAATTCATAAGCTTATACTAGTGTC*TTAATGTTCAACCTTAACTTCTCC
oJZ259	*CATGCATTAGCTAGCATTAC*AGGAGGCTAGCCTATG
oJZ394	*GAAATACATAAAAGATAATATAGGCTGGAA*AGG
oJZ395	*GGAGTTCAAACTTGATATCGAATTCCTG*
oJZ396	*CAGGAATTCGATATCAAGTTTGAACTCC*TTTTTCATATGAGAAG
oJZ397	*TTTTCTGACTCATAGTAGTTCCTCCTTC*TAGATTTTTTTTTGAATTCTACAG
oJZ398	*AAGGAGGAACTACTATGAGTCAGAAAA*CAGACG
oJZ399	*TTCCAGCCTATATTATCTTTTATGTATTTC*TAGCTTCAGAG
oJZ436	CTGGGGGCTGGGATACTTTG
oJZ437	TCTCCCAAGCCAGAACCAAC

* Restriction sites underlined, Gibson assembly or overlap extension PCR homology flanks italicized

Supplementary Table 7

Strains used in Appendix A

Strain	Genotype	Parent strain	Reference
BL21(DE3)pLysS	F(-) ompT gal dcm lon hsdSB(rB(-) mB(-)) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) ; pLysS	N/A	(184)
eJZ110	pJZ085	BL21(DE3)pLysS	This study

Supplementary Table 8

Plasmids used in Appendix A

Plasmid	Description	Source
pJZ085	*P_{T7}-6XHis-SUMO-comGC$^{\Delta 1\text{-}30}$ (kan, cm)*	GenScript

Supplementary strain construction

Chapter 2

Note: All strains below are isogenic with PY79, unless otherwise listed.

bJZ096 [*lacA::P_{comG}-comGC^{E56C} (erm)*] was generated by first amplifying a pair of PCR products using primer pairs (**U-F** and **C-F-56**) and (**U-R** and **C-R-56**) according to the SOE PCR protocol described above. These primer pairs were stitched together, and then amplified, following the same protocol. The amplified product was directly moved into **PY79** via natural transformation as described above.

bJZ097 [*lacA::P_{comG}-comGC^{L57C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-57**) and (**U-R** and **C-R-57**).

bJZ098 [*lacA::P_{comG}-comGC^{D58C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-58**) and (**U-R** and **C-R-58**).

bJZ099 [*lacA::P_{comG}-comGC^{H59C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-59**) and (**U-R** and **C-R-59**).

bJZ100 [*lacA::P_{comG}-comGC^{G61C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-61**) and (**U-R** and **C-R-61**).

bJZ101 [*lacA::P_{comG}-comGC^{Q62C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-62**) and (**U-R** and **C-R-62**).

bJZ102 [*lacA::P_{comG}-comGC^{T63C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-63**) and (**U-R** and **C-R-63**).

bJZ103 [*lacA::P_{comG}-comGC^{S65C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-65**) and (**U-R** and **C-R-65**).

bJZ104 [*lacA::P_{comG}-comGC^{L66C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-66**) and (**U-R** and **C-R-66**).

bJZ105 [*lacA::P_{comG}-comGC^{A67C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-67**) and (**U-R** and **C-R-67**).

bJZ106 [*lacA::P_{comG}-comGC^{D68C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-68**) and (**U-R** and **C-R-68**).

bJZ107 [*lacA::P_{comG}-comGC^{L69C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-69**) and (**U-R** and **C-R-69**).

bJZ108 [*lacA::P_{comG}-comGC^{Q70C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-70**) and (**U-R** and **C-R-70**).

bJZ109 [*lacA::P_{comG}-comGC^{S71C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-71**) and (**U-R** and **C-R-71**).

bJZ110 [*lacA::P_{comG}-comGC^{E72C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-72**) and (**U-R** and **C-R-72**).

bJZ111 [*lacA::P_{comG}-comGC^{G73C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-73**) and (**U-R** and **C-R-73**).

bJZ112 [*lacA::P_{comG}-comGC^{Y74C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-74**) and (**U-R** and **C-R-74**).

bJZ113 [*lacA::*P_{comG}*-comGC*^{K76C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-76**) and (**U-R** and **C-R-76**).

bJZ114 [*lacA::*P_{comG}*-comGC*^{K77C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-77**) and (**U-R** and **C-R-77**).

bJZ115 [*lacA::*P_{comG}*-comGC*^{D78C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-78**) and (**U-R** and **C-R-78**).

bJZ116 [*lacA::*P_{comG}*-comGC*^{V80C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-80**) and (**U-R** and **C-R-80**).

bJZ117 [*lacA::*P_{comG}*-comGC*^{C81C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-81**) and (**U-R** and **C-R-81**).

bJZ118 [*lacA::*P_{comG}*-comGC*^{N83C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-83**) and (**U-R** and **C-R-83**).

bJZ119 [*lacA::*P_{comG}*-comGC*^{G84C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-84**) and (**U-R** and **C-R-84**).

bJZ120 [*lacA::*P_{comG}*-comGC*^{K85C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-85**) and (**U-R** and **C-R-85**).

bJZ121 [*lacA::*P_{comG}*-comGC*^{R86C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-86**) and (**U-R** and **C-R-86**).

bJZ122 [*lacA::*P_{comG}*-comGC*^{I88C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-88**) and (**U-R** and **C-R-88**).

bJZ123 [*lacA::*P_{comG}*-comGC*^{I89C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-89**) and (**U-R** and **C-R-89**).

bJZ124 [*lacA::*P_{comG}*-comGC*^{G92C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-92**) and (**U-R** and **C-R-92**).

bJZ125 [*lacA::*P_{comG}*-comGC*^{E93C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-93**) and (**U-R** and **C-R-93**).

bJZ126 [*lacA::*P_{comG}*-comGC*^{V94C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-94**) and (**U-R** and **C-R-94**).

bJZ127 [*lacA::*P_{comG}*-comGC*^{K95C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-95**) and (**U-R** and **C-R-95**).

bJZ128 [*lacA::*P_{comG}*-comGC*^{V96C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-96**) and (**U-R** and **C-R-96**).

bJZ129 [*lacA::*P_{comG}*-comGC*^{E97C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-97**) and (**U-R** and **C-R-97**).

bJZ130 [*lacA::*P_{comG}*-comGC*^{H98C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-98**) and (**U-R** and **C-R-98**).

bJZ133 [*lacA::*P_{comG}*-comGC*^{E60C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-60**) and (**U-R** and **C-R-60**).

bJZ134 [*lacA::*P_{comG}*-comGC*^{P64C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-64**) and (**U-R** and **C-R-64**).

bJZ135 [*lacA::*P_{comG}*-comGC*^{V75C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-75**) and (**U-R** and **C-R-75**).

bJZ136 [*lacA::*P_{comG}*-comGC*^{P82C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-82**) and (**U-R** and **C-R-82**).

bJZ137 [*lacA::*P_{comG}*-comGC*^{I87C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-87**) and (**U-R** and **C-R-87**).

bJZ138 [*lacA::*P_{comG}*-comGC*^{T90C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-90**) and (**U-R** and **C-R-90**).

bJZ139 [*lacA::*P_{comG}*-comGC*^{G91C} *(erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-91**) and (**U-R** and **C-R-91**).

bJZ140 [*lacA::*P_{comG}*-comGC*^{E56C} *(erm) ; comGC::kan*] was generated by natural transformation of **bJZ096** using a PCR product (*comGC::kan*) obtained by amplifying from **BKK24710** chromosomal DNA according to the PCR protocol for validating *comGC* deletions described above. Putative transformants were screened for resistance to both MLS and Kan[10] before advancing. Integration of P_{comG}*-comGC*^{E56C} *(erm)* into the *lacA* locus was confirmed by the PCR validation procedure noted above, and the sequence was confirmed by Sanger sequencing of the PCR product made during the validation procedure (using sequencing primer **oJZ153**). Deletion of *comGC* was confirmed according to the PCR validation protocol described above.

bJZ141 [*lacA::*P_{comG}*-comGC*^{D58C} *(erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ098**

bJZ142 [*lacA::*P_{comG}*-comGC*^{H59C} *(erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ099**

bJZ143 [*lacA::*P_{comG}*-comGC*^{G61C} *(erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ100**

bJZ144 [*lacA::*P_{comG}*-comGC*^{Q62C} *(erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ101**

bJZ145 [*lacA::*P_{comG}*-comGC*^{T63C} *(erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ102**

bJZ146 [*lacA::*P_{comG}*-comGC*^{S65C} *(erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ103**

bJZ147 [*lacA::*P_{comG}*-comGC*^{A67C} *(erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ105**

bJZ148 [*lacA::*P_{comG}*-comGC*^{L69C} *(erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ107**

bJZ149 [*lacA::*P_{comG}*-comGC*^{Q70C} *(erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ108**

bJZ150 [*lacA::*P_{comG}*-comGC*^{S71C} *(erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ109**

bJZ151 [*lacA::*P_{comG}*-comGC*^{E72C} *(erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ110**

bJZ152 [*lacA::P$_{comG}$-comGC^{G73C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ111**

bJZ153 [*lacA::P$_{comG}$-comGC^{K76C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ113**

bJZ154 [*lacA::P$_{comG}$-comGC^{K77C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ114**

bJZ155 [*lacA::P$_{comG}$-comGC^{V80C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ116**

bJZ156 [*lacA::P$_{comG}$-comGC^{C81C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ117**

bJZ157 [*lacA::P$_{comG}$-comGC^{N83C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ118**

bJZ158 [*lacA::P$_{comG}$-comGC^{G84C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ119**

bJZ159 [*lacA::P$_{comG}$-comGC^{K85C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ120**

bJZ160 [*lacA::P$_{comG}$-comGC^{R86C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ121**

bJZ161 [*lacA::P$_{comG}$-comGC^{I88C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ122**

bJZ162 [*lacA::P$_{comG}$-comGC^{I89C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ123**

bJZ163 [*lacA::P$_{comG}$-comGC^{K95C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ127**

bJZ164 [*lacA::P$_{comG}$-comGC^{E97C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ129**

bJZ165 [*lacA::P$_{comG}$-comGC^{H98C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ130**

bJZ166 [*lacA::P$_{comG}$-comGC^{P64C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ134**

bJZ167 [*lacA::P$_{comG}$-comGC^{P82C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ136**

bJZ168 [*lacA::P$_{comG}$-comGC^{I87C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ137**

bJZ169 [*lacA::P$_{comG}$-comGC^{T90C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ138**

bJZ170 [*lacA::P$_{comG}$-comGC^{G91C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ139**

bJZ171 [*lacA::P_{comG}-comGC^{A79C} (erm)*] was generated using the same protocol for **bJZ096**, but initial PCR used primer pairs (**U-F** and **C-F-79**) and (**U-R** and **C-R-79**).

bJZ172 [*lacA::P_{comG}-comGC^{L57C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ097**

bJZ173 [*lacA::P_{comG}-comGC^{D68C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ106**

bJZ174 [*lacA::P_{comG}-comGC^{Y74C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ112**

bJZ175 [*lacA::P_{comG}-comGC^{D78C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ115**

bJZ176 [*lacA::P_{comG}-comGC^{G92C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ124**

bJZ177 [*lacA::P_{comG}-comGC^{E93C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ125**

bJZ178 [*lacA::P_{comG}-comGC^{V94C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ126**

bJZ179 [*lacA::P_{comG}-comGC^{V96C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ128**

bJZ180 [*lacA::P_{comG}-comGC^{E60C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ133**

bJZ181 [*lacA::P_{comG}-comGC^{V75C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ135**

bJZ182 [*lacA::P_{comG}-comGC^{A79C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ171**

bJZ183 [*lacA::P_{comG}-comGC^{L66C} (erm) ; comGC::kan*] was generated/validated/sequenced by using the same protocol for **bJZ140**, but the (*comGC::kan*) PCR product was naturally transformed into **bJZ104**

bJZ238 [*comGC::kan*] was generated by natural transformation of **PY79** using a PCR product (*comGC::kan*) obtained by amplifying from **BKK24710** chromosomal DNA according to the PCR protocol for validating *comGC* deletions described above.

bJZ363 [*amyE::$P_{spank\text{-}hy}$-comG (spc)*] was generated by natural transformation of SpeI-linearized **pBOSE1620** into **PY79.**

bJZ447 [*amyE::$P_{spank\text{-}hy}$ rapI-catR ; ICEBs1 Δ(rapI-phrI)::kan ; ΔattR::tet ; ΔnicK(unmarked) (oriT-) ; pDR244*] was generated by natural transformation of intact **pDR244** into **ELC453**. Plasmid maintained in strain by culturing with Spc[50].

bJZ450 [*lacA::P_{comG}-comGC^{E56C} (erm) ; comGC markerless deletion*] was generated by the Cre-Lox mediated recombination of *comGC::kan* in **bJZ140** using the *kan* cassette removal protocol described above.

bJZ451 [*lacA::P_{comG}-comGC^{D58C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ141**

bJZ452 [*lacA::P_{comG}-comGC^{H59C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ142**

bJZ453 [*lacA::P_{comG}-comGC^{G61C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ143**

bJZ454 [*lacA::P_{comG}-comGC^{Q62C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ144**

bJZ455 [*lacA::P_{comG}-comGC^{T63C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ145**

bJZ456 [*lacA::P_{comG}-comGC^{S65C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ146**

bJZ457 [*lacA::P_{comG}-comGC^{A67C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ147**

bJZ458 [*lacA::P_{comG}-comGC^{L69C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ148**

bJZ459 [*lacA::P_{comG}-comGC^{Q70C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ149**

bJZ460 [*lacA::P_{comG}-comGC^{S71C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ150**

bJZ461 [*lacA::P_{comG}-comGC^{E72C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ151**

bJZ462 [*lacA::P_{comG}-comGC^{G73C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ152**

bJZ463 [*lacA::P_{comG}-comGC^{K76C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ153**

bJZ464 [*lacA::P_{comG}-comGC^{K77C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ154**

bJZ465 [*lacA::P_{comG}-comGC^{V80C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ155**

bJZ466 [*lacA::P_{comG}-comGC^{C81C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ156**

bJZ467 [*lacA::P_{comG}-comGC^{N83C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ157**

bJZ468 [*lacA::P_{comG}-comGC^{G84C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ158**

bJZ469 [*lacA::P_{comG}-comGC^{K85C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ159**

bJZ470 [*lacA::P_{comG}-comGC^{R86C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ160**

bJZ471 [*lacA::P_{comG}-comGC^{I88C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ161**

bJZ472 [*lacA::P_{comG}-comGC^{I89C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ162**

bJZ473 [*lacA::P_{comG}-comGC^{K95C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ163**

bJZ474 [*lacA::P_{comG}-comGC^{E97C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ164**

bJZ475 [*lacA::P_{comG}-comGC^{H98C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ165**

bJZ476 [*lacA::P_{comG}-comGC^{P64C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ166**

bJZ477 [*lacA::P_{comG}-comGC^{P82C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ167**

bJZ478 [*lacA::P_{comG}-comGC^{I87C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ168**

bJZ479 [*lacA::P_{comG}-comGC^{T90C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ169**

bJZ480 [*lacA::P_{comG}-comGC^{G91C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ170**

bJZ481 [*lacA::P_{comG}-comGC^{L57C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ172**

bJZ482 [*lacA::P_{comG}-comGC^{D68C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ173**

bJZ483 [*lacA::P_{comG}-comGC^{Y74C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ174**

bJZ484 [*lacA::P_{comG}-comGC^{D78C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ175**

bJZ485 [*lacA::P_{comG}-comGC^{G92C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ176**

bJZ486 [*lacA::P_{comG}-comGC^{E93C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ177**

bJZ487 [*lacA::P_{comG}-comGC^{V94C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ178**

bJZ488 [*lacA::P_{comG}-comGC^{V96C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ179**

bJZ489 [*lacA::P_{comG}-comGC^{E60C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ180**

bJZ490 [*lacA::P_{comG}-comGC^{V75C} (erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ181**

bJZ491 [*lacA::*P_{comG}*-comGC*^{A79C} *(erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ182**

bJZ492 [*lacA::*P_{comG}*-comGC*^{L66C} *(erm) ; comGC markerless deletion*] was generated by repeating the **bJZ450** construction protocol using **bJZ183**

Chapter 3

Note: All strains below are isogenic with PY79, unless otherwise listed.

eJZ040 [pJZ026 (*yhdGH::$P_{xyl(-CRE)}$ (kan)*)] was generated by transformation of Gibson assembled **pJZ026** into chemically competent **DH5α** via heat-shock, selecting for Amp^R (157). Transformants were verified by colony PCR screening.

eJZ064 [pJZ049 (*yhdGH::$P_{xyl(-CRE)}$-rbs-comGC (kan)*)] was generated by transformation of Gibson assembled **pJZ049** into chemically competent **DH5α** via heat-shock, selecting for Amp^R (157). Transformants were verified by colony PCR screening.

eJZ095 [pJZ069 (*yhdGH::P_{xyl}-rbs-comK (kan)*)] was generated by transformation of Gibson assembled **pJZ069** into chemically competent **DH5α** via heat-shock, selecting for Amp^R (157). Transformants were verified by colony PCR screening.

bBB036 [*comK::spc*] was generated by transformation of BD2259 (185) genomic DNA into competent **PY79** selecting Spc^R, followed by two successive backcrosses of transformant genomic DNA into competent **PY79** selecting Spc^R.

bBB050 [*amyE::PcomE-gfp-comEA (cat)*] was generated by a double crossover of **pBB401** into competent **PY79** selecting Cm^R. The double crossover was confirmed by PCR from genomic DNA.

bJZ096 [*lacA::P_{comG}-comGC^{E56C} (erm)*] was generated first by overlap extension PCR (186) that produced *lacA::P_{comG}-comGC^{E56C} (erm)*. Primers pairs (**oJZ149** + **oJZ152**) and (**oJZ150** + **oJZ151**) were used in conventional PCR with **pKRH83** to generate two *lacA::P_{comG}-comGC^{E56C} (erm)* fragments with two homologous ends. The products of each reaction were isolated and extracted from an agarose gel (157), and an equimolar ratio of each purified product was used in a subsequent PCR (without primers) to make full length *lacA::P_{comG}-comGC^{E56C} (erm)*. Primer pair (**oJZ149** + **oJZ150**) was added approximately halfway through the reaction to begin exponential amplification of full length *lacA::P_{comG}-comGC^{E56C} (erm)*. Once amplification was verified by agarose gel electrophoresis, full length *lacA::P_{comG}-comGC^{E56C} (erm)* PCR product was used to directly transform competent **PY79** selecting MLS^R. Double crossover recombination was verified by PCR of genomic DNA using primer pair (**oJZ212** + **oJZ213**) and the same PCR product was Sanger sequenced (primer **oJZ214**) to verify that no mutations other than *comGC^{E56C}* had been introduced.

bJZ103 [*lacA::P_{comG}-comGC^{S65C} (erm)*] was generated as described for **bJZ096**, but the conventional PCR primer pairs were swapped to (**oJZ149** + **oJZ176**) and (**oJZ150** + **oJZ175**).

bJZ117 [*lacA::P_{comG}-comGC^{WT} (erm)*] was generated as described for **bJZ096**, but the conventional PCR primer pairs were swapped to (**oJZ149** + **oJZ202**) and (**oJZ150** + **oJZ201**).

bJZ140 [*lacA::P_{comG}-comGC^{E56C} (erm) ; comGC::kan*] was generated by transformation of competent **bJZ096** with *comGC::kan* PCR product (primer pair **oJZ223** + **oJZ225** with **bJZ083** genomic DNA template) selecting Kan^R followed by screening 10 transformants for MLS^R by patching single colonies. Double crossover recombination of comGC::kan was confirmed by PCR using primer pair (**oJZ199** + **oJZ200**) and strain genomic DNA.

bJZ146 [*lacA::P_{comG}-comGC^{S65C} (erm) ; comGC::kan*] was generated as described for **bJZ140**, but competent **bJZ103** was transformed with *comGC::kan* PCR product

bJZ156 [*lacA::P_{comG}-comGC^{WT} (erm) ; comGC::kan*] was generated as described for **bJZ140**, but competent **bJZ117** was transformed with *comGC::kan* PCR product

bJZ405 [*yhdGH::P_{xyl}-rbs-comK (kan)*] was generated first by preparing **pJZ069** plasmid DNA from **eJZ095** by alkaline lysis (157). **pJZ069** was then restriction digested with ScaI-HF to linearize (to facilitate double crossover recombination), and then transformed into competent **PY79** selecting for Kan^R. The transformed allele was tested for functionality by complementation in **bJZ407**.

bJZ407 [*yhdGH::P_{xyl}-rbs-comK (kan) ; comK::spc*] was generated by transforming competent **bJZ405** with genomic DNA from **bBB036** and selecting for Spc^R followed by screening 10 transformants for Kan^R by patching single colonies. Kan^R Spc^R transformants were tested for *comK* complementation by transformation of induced (0.5% (w/v) xylose) and non-induced cultures with lysed **bBB050** protoplasts (genomic DNA carrying *amyE::PcomE-gfp-comEA (cat)*) and selecting for Cm^R. All transformants tested were non-transformable when not induced, and fully transformable when induced, indicating full complementation of *comK::spc*.

bJZ418 [*yhdGH::P_{xyl}-rbs-comK (kan) ; amyE::P_{comE}-gfp-comEA (cat)*] was generated by transforming competent **bJZ405** with genomic DNA from **bBB050** selecting for Cm^R. Transformants were expected to express *gfp-comEA* in a much greater proportion of cells when induced with 0.5% xylose, and this was confirmed by epifluorescence microscopy.

bJZ419 [*yhdGH::P_{xyl}-rbs-comK (kan) ; lacA::P_{comG}-comGC^{E56C} (erm)*] was generated by transforming competent **bJZ405** with genomic DNA from **bJZ096** selecting for MLS^R. Double crossover recombination was verified by PCR with primer pair (oBOSE397 + oBOSE400) with strain genomic DNA as template.

bJZ423 [*yhdGH::P_{xyl}-rbs-comK (kan) ; lacA::P_{comG}-comGC^{S65C} (erm)*] was generated as described for **bJZ419**, but genomic DNA from **bJZ103** was used in the transformation.

bJZ427 [*yhdGH::P_{xyl}-rbs-comK (kan) ; lacA::P_{comG}-comGC^{WT} (erm)*] was generated as described for **bJZ419**, but genomic DNA from **bJZ117** was used in the transformation.

Appendix A

Note: All strains below are isogenic with PY79, unless otherwise listed.

eJZ110 [pJZ085] was generated by heat shock transformation of chemically competent BL21(DE3)pLysS with pJZ085.

Supplementary plasmid construction

Chapter 3

pBB401 [*amyE::PcomE-gfp-comEA(cat)*] was generated in a four-way ligation with an *Eco*RI-*Hind*III PCR product containing the promoter region for the *comE* operon (oligonucleotide primers oBMB116 and oBMB117 and genomic DNA from **PY79** as template), an *Hind*III-*Xho*I PCR product containing an optimized RBS and *gfp* (oligonucleotides oBMB127 and oBMB128 and **pKL147** as template), an *Xho*I-*Bam*HI PCR product containing *comEA*, all cloned into **pNS008** cut with *Eco*RI and *Bam*HI. **pNS008** was a gift from D. Rudner.

pJZ026 [*yhdGH::*$P_{xyl(-CRE)}$ *(kan)*] was generated by a 2-fragment Gibson assembly. Primer pair 1 (**oJZ167** + **oJZ168**) and primer pair 2 (**oJZ169** + **oJZ170**) were used in separate PCRs with **pBOSE1408** as template in each reaction to amplify two fragments of **pBOSE1408** comprising almost the entire **pBOSE1408** sequence but lacking a catabolite repressive element found in the original P_{xyl} cassette (187). Each PCR product was isolated and extracted from an agarose gel, and an equimolar ratio of each fragment was used to construct **pJZ026** by a 2-fragment Gibson assembly using HiFi master mix (NEB) according to the manufacturer's instructions (60 min incubation time).

pJZ049 [*yhdGH::*$P_{xyl(-CRE)}$*-rbs-comGC (kan)*] was generated by a 2-fragment Gibson assembly. **pJZ026** was restriction digested with XhoI (NEB) to linearize for Gibson assembly. Primer pair 1 (**oJZ257** + **oJZ258**) was used in PCR with **pKRH83** as template to amplify *rbs-comGC* (**pKRH83** was a gift from D. Kearns). The amplified *rbs-comGC* product was extracted from an agarose gel and used as template in a subsequent PCR with primer pair 2 (**oJZ258** +**oJZ259**) to amplify *rbs-comGC* with appropriate homologous ends for Gibson assembly. XhoI-digested **pJZ026** and the subsequent *rbs-comGC* amplified product were extracted from an agarose gel, and an equimolar ratio of each fragment was used to construct **pJZ049** by a 2-fragment Gibson assembly using HiFi master mix (NEB) according to the manufacturer's instructions (60 min incubation time).

pJZ069 [*yhdGH::*P_{xyl}*-rbs-comK (kan)*] was generated by a 3-fragment Gibson assembly. Primer pair 1 (**oJZ394** + **oJZ395**) was used in PCR with **pJZ049** to amplify the vector backbone and *yhdGH::kan* ectopic integration sequence, primer pair 2 (**oJZ396** + **oJZ397**) was used in PCR with **pBS0E-XylR-**$\mathbf{P_{xylA\ (V2)}}$ (145) to amplify the P_{xyl} *(xylR)* cassette, and primer pair 3 (**oJZ398** + **oJZ399**) was used in PCR with **PY79** genomic DNA to amplify *comK* including its transcriptional terminator. Each PCR product was isolated and extracted from an agarose gel, and an equimolar ratio of each fragment was used to construct **pJZ069** by a 3-fragment Gibson assembly using HiFi master mix (NEB) according to the manufacturer's instructions (60 min incubation time). The entire plasmid was then sequenced by Plasmidsaurus Oxford Nanopore sequencing for confirmation.

Appendix A

pJZ085 [P_{T7}-*6XHis-SUMO-comGC*$^{\Delta 1\text{-}30}$ *(kan, cm)*] was generated by GenScript through subcloning of a synthesized gene encoding *comGC*$^{\Delta 1\text{-}30}$ into the pE-SUMOpro (Kan ; LifeSensors) vector digested with NcoI and EcoRI restriction enzymes. (188)

References

1. Gregory TR. 2009. Understanding Natural Selection: Essential Concepts and Common Misconceptions. Evolution: Education and Outreach 2:156–175.

2. Lenski RE. 2017. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. The ISME Journal 2017 11:10 11:2181–2194.

3. Blount ZD, Borland CZ, Lenski RE. 2008. Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. Proc Natl Acad Sci U S A 105:7899–7906.

4. Abe K, Nomura N, Suzuki S. 2020. Biofilms: hot spots of horizontal gene transfer (HGT) in aquatic environments, with a focus on a new HGT mechanism. FEMS Microbiol Ecol 96:31.

5. Arnold BJ, Huang IT, Hanage WP. 2021. Horizontal gene transfer and adaptive evolution in bacteria. Nature Reviews Microbiology 2021 20:4 20:206–218.

6. Chiang YN, Penadés JR, Chen J. 2019. Genetic transduction by phages and chromosomal islands: The new and noncanonical. PLoS Pathog 15.

7. Saragliadis A, Trunk T, Leo JC. 2018. Producing Gene Deletions in Escherichia coli by P1 Transduction with Excisable Antibiotic Resistance Cassettes. J Vis Exp 2018:58267.

8. Cabezón E, Ripoll-Rozada J, Peña A, de la Cruz F, Arechaga I. 2015. Towards an integrated model of bacterial conjugation. FEMS Microbiol Rev 39:81–95.

9. Dubey GP, Ben-Yehuda S. 2011. Intercellular nanotubes mediate bacterial communication. Cell 144:590–600.

10. Johnston C, Martin B, Fichant G, Polard P, Claverys JP. 2014. Bacterial transformation: distribution, shared mechanisms and divergent control. Nature Reviews Microbiology 2014 12:3 12:181–196.

11. Berka RM, Hahn J, Albano M, Draskovic I, Persuh M, Cui X, Sloma A, Widner W, Dubnau D. 2002. Microarray analysis of the Bacillus subtilis K-state: genome-wide expression changes dependent on ComK. Mol Microbiol 43:1331–1345.

12. Chen I, Dubnau D. 2004. DNA uptake during bacterial transformation. Nature Reviews Microbiology 2004 2:3 2:241–249.

13. Dubnau D. 1999. DNA uptake in bacteria. Annu Rev Microbiol 53:217–244.

14. Claverys JP, Prudhomme M, Martin B. 2006. Induction of competence regulons as a general response to stress in gram-positive bacteria. Annu Rev Microbiol 60:451–475.

15. Mongold JA. 1992. DNA Repair and the Evolution of Transformation in Haemophilus Influenzae. Genetics 132:893.

16. Mell JC, Redfield RJ. 2014. Natural competence and the evolution of DNA uptake specificity. J Bacteriol 196:1471–1483.

17. Vogels GD, Drift C Van der. 1976. Degradation of purines and pyrimidines by microorganisms. Bacteriol Rev 40:403.

18. Hall RJ, Whelan FJ, McInerney JO, Ou Y, Domingo-Sananes MR. 2020. Horizontal Gene Transfer as a Source of Conflict and Cooperation in Prokaryotes. Front Microbiol 11:536378.

19. Maier B. 2020. Competence and Transformation in Bacillus subtilis. Curr Issues Mol Biol 37:57–76.

20. Hamilton L, Dillard JP, Hamilton HL, Dillard JP. 2006. Natural transformation of Neisseria gonorrhoeae: from DNA donation to homologous recombination. Mol Microbiol 59:376–385.

21. Matthey N, Blokesch M. 2016. The DNA-Uptake Process of Naturally Competent Vibrio cholerae https://doi.org/10.1016/j.tim.2015.10.008.

22. Salvadori G, Junges R, Morrison DA, Petersen FC. 2019. Competence in streptococcus pneumoniae and close commensal relatives: Mechanisms and implications. Front Cell Infect Microbiol 9:451086.

23. Su Y, Liu C, Fang H, Zhang D. 2020. Bacillus subtilis: a universal cell factory for industry, agriculture, biomaterials and medicine. Microbial Cell Factories 2020 19:1 19:1–12.

24. Liu Y, Zhang Q, Qi X, Gao H, Wang M, Guan H, Yu B. 2023. Metabolic Engineering of Bacillus subtilis for Riboflavin Production: A Review. Microorganisms 11.

25. Li Y, Li G, Zhao X, Shao Y, Wu M, Ma T. 2019. Regulation of hyaluronic acid molecular weight and titer by temperature in engineered Bacillus subtilis. 3 Biotech 9:1–9.

26. Yang H, Ma Y, Zhao Y, Shen W, Chen X. 2020. Systematic engineering of transport and transcription to boost alkaline α-amylase production in Bacillus subtilis. Appl Microbiol Biotechnol 104:2973–2985.

27. Godeux AS, Svedholm E, Barreto S, Potron A, Venner S, Charpentier X, Laaberki MH. 2022. Interbacterial Transfer of Carbapenem Resistance and Large Antibiotic Resistance Islands by Natural Transformation in Pathogenic Acinetobacter. mBio 13.

28. Domingues S, Rosário N, Cândido Â, Neto D, Nielsen KM, Da Silva GJ. 2019. Competence for Natural Transformation Is Common among Clinical Strains of Resistant Acinetobacter spp. Microorganisms 7.

29. Strahl H, Errington J. 2017. Bacterial Membranes: Structure, Domains, and Function. Annu Rev Microbiol 71:519–538.

30. Zhang YM, Rock CO. 2008. Membrane lipid homeostasis in bacteria. Nat Rev Microbiol 6:222–233.

31. Chang S, Cohen SN. 1979. High frequency transformation of Bacillus subtilis protoplasts by plasmid DNA. MGG Molecular & General Genetics 168:111–115.

32. Nikaido H. 2003. Molecular Basis of Bacterial Outer Membrane Permeability Revisited. Microbiology and Molecular Biology Reviews 67:593–656.

33. Silhavy TJ, Kahne D, Walker S. 2010. The Bacterial Cell Envelope. Cold Spring Harb Perspect Biol 2.

34. Yao X, Jericho M, Pink D, Beveridge T. 1999. Thickness and elasticity of gram-negative murein sacculi measured by atomic force microscopy. J Bacteriol 181:6865–6875.

35. Vollmer W, Höltje JV. 2004. The architecture of the murein (peptidoglycan) in gram-negative bacteria: Vertical scaffold or horizontal layer(s)? J Bacteriol. American Society for Microbiology Journals https://doi.org/10.1128/JB.186.18.5978-5987.2004.

36. Matias VRF, Beveridge TJ. 2008. Lipoteichoic acid is a major component of the Bacillus subtilis periplasm. J Bacteriol 190:7414–7418.

37. Malanovic N, Lohner K. 2016. Gram-positive bacterial cell envelopes: The impact on the activity of antimicrobial peptides. Biochim Biophys Acta Biomembr 1858:936–946.

38. Brown S, Santa Maria JP, Walker S. 2013. Wall teichoic acids of gram-positive bacteria. Annu Rev Microbiol 67:313–336.

39. Dickson JS, Koohmaraie M, Hruska RL. 1989. Cell Surface Charge Characteristics and Their Relationship to Bacterial Attachment to Meat SurfacesAPPLIED AND ENVIRONMENTAL MICROBIOLOGY.

40. Griffith F. 1928. The Significance of Pneumococcal Types. J Hyg (Lond) 27:113.

41. Spizizen J. 1958. TRANSFORMATION OF BIOCHEMICALLY DEFICIENT STRAINS OF BACILLUS SUBTILIS BY DEOXYRIBONUCLEATE. Proc Natl Acad Sci U S A 44:1072.

42. Anagnostopoulos C, Spizizen J. 1961. REQUIREMENTS FOR TRANSFORMATION IN BACILLUS SUBTILIS. J Bacteriol 81:741.

43. Young FE, Spizizen J. 1963. INCORPORATION OF DEOXYRIBONUCLEIC ACID IN THE BACILLUS SUBTILIS TRANSFORMATION SYSTEM. J Bacteriol 86:392.

44. YOUNG FE, SPIZIZEN J. 1961. PHYSIOLOGICAL AND GENETIC FACTORS AFFECTING TRANSFORMATION OF BACILLUS SUBTILIS. J Bacteriol 81:823.

45. Fox MS. 1957. Deoxyribonucleic acid incorporation by transformed bacteria. Biochim Biophys Acta 26:83–85.

46. Maurice DR, Fox S, Rollin DR, Hotchkiss D. 1957. Initiation of Bacterial Transformation. Nature 1957 179:4574 179:1322–1325.

47. Fox MS, Hotchkiss RD. 1960. Fate of Transforming Deoxyribonucleate following Fixation by Transformable Bacteria: I. Nature 1960 187:4742 187:1002–1003.

48. ALEXANDER HE, LEIDY G. 1953. Induction of streptomycin resistance in sensitive Hemophilus influenzae by extracts containing desoxyribonucleic acid from resistant Hemophilus influenzae. J Exp Med 97:17–31.

49. ALEXANDER HE, LEIDY G, HAHN E. 1954. Studies on the nature of hemophilus influenzae cells susceptible to heritable changes by desoxyribonucleic acids. J Exp Med 99:505–533.

50. GOODGAL SH, HERRIOTT RM. 1961. Studies on transformations of Hemophilus influenzae. I. Competence. J Gen Physiol 44:1201–1227.

51. STUY JH, STERN D. 1964. THE KINETICS OF DNA UPTAKE BY HAEMOPHILUS INFLUENZAE. J Gen Microbiol 35:391–400.

52. CATLIN BW. 1960. TRANSFORMATION OF NEISSERIA MENINGITIDIS BY DEOXYRIBONUCLEATES FROM CELLS AND FROM CULTURE SLIME. J Bacteriol 79:579.

53. LIE S. 1965. STUDIES ON THE PHENOTYPIC EXPRESSION OF COMPETENCE IN NEISSERIA MENINGITIDIS. Acta Pathol Microbiol Scand 64:119–129.

54. Lacks S. 1962. Molecular fate of DNA in genetic transformation of Pneumococcus. J Mol Biol 5:119–131.

55. Barnhart BJ, Herriott RM. 1963. Penetration of deoxyribonucleic acid into Hemophilus influenzae. Biochimica et Biophysica Acta (BBA) - Specialized Section on Nucleic Acids and Related Subjects 76:25–39.

56. Bodmer WF, Ganesan AT. 1964. BIOCHEMICAL AND GENETIC STUDIES OF INTEGRATION AND RECOMBINATION IN BACILLUS SUBTILIS TRANSFORMATION. Genetics 50:717–738.

57. Dubnau D, Davidoff-Abelson R. 1971. Fate of transforming DNA following uptake by competent Bacillus subtilis: I. Formation and properties of the donor-recipient complex. J Mol Biol 56:209–221.

58. Davidoff-Abelson R, Dubnau D. 1971. Fate of transforming DNA after uptake by competent Bacillus subtilis: failure of donor DNA to replicate in a recombination-deficient recipient. Proc Natl Acad Sci U S A 68:1070–1074.

59. Dubnau D, Cirigliano C. 1972. Fate of transforming DNA following uptake by competent Bacillus subtilis: III. Formation and properties of products isolated from transformed cells which are derived entirely from donor DNA. J Mol Biol 64:9–29.

60. Dubnau D, Cirigliano C. 1972. Fate of transforming DNA following uptake by competent Bacillus subtilis. IV. The endwise attachment and uptake of transforming DNA. J Mol Biol 64:31–46.

61. FOX MS, ALLEN MK. 1964. ON THE MECHANISM OF DEOXYRIBONUCLEATE INTEGRATION IN PNEUMOCOCCAL TRANSFORMATION. Proc Natl Acad Sci U S A 52:412.

62. Lacks S, Greenberg B, Carlson K. 1967. Fate of donor DNA in pneumococcal transformation. J Mol Biol 29:327–347.

63. Morrison DA, Guild WR. 1973. Structure of deoxyribonucleic acid on the cell surface during uptake by pneumococcus. J Bacteriol 115:1055–1062.

64. Notani N, Goodgal SH. 1966. On the nature of recombinants formed during transformation in Hemophilus influenzae. J Gen Physiol 49:197–209.

65. Goodgal SH, Notani N. 1968. Evidence that either strand of DNA can transform. J Mol Biol 35:449–453.

66. Lacks S, Greenberg B, Neuberger M. 1975. Identification of a deoxyribonuclease implicated in genetic transformation of Diplococcus pneumoniae. J Bacteriol 123:222–232.

67. Lacks S, Neuberger M. 1975. Membrane location of a deoxyribonuclease implicated in the genetic transformation of Diplococcus pneumoniae. J Bacteriol 124:1321–1329.

68. Lacks S. 1970. Mutants of Diplococcus pneumoniae that lack deoxyribonucleases and other activities possibly pertinent to genetic transformation. J Bacteriol 101:373–383.

69. Fani R, Mastromei G, Polsinelli M, Venema G. 1984. Isolation and characterization of Bacillus subtilis mutants altered in competence. J Bacteriol 157:152–157.

70. Mulder JA, Venema G. 1982. Isolation and partial characterization of Bacillus subtilis mutants impaired in DNA entry. J Bacteriol 150:260–268.

71. Mulder JA, Venema G. 1982. Transformation-deficient mutants of Bacillus subtilis impaired in competence-specific nuclease activities. J Bacteriol 152:166.

72. Smith H, De Vos W, Bron S. 1983. Transformation in Bacillus subtilis: properties of DNA-binding-deficient mutants. J Bacteriol 153:12–20.

73. Smith H, Wiersma K, Venema G, Bron S. 1985. Transformation in Bacillus subtilis: further characterization of a 75,000-dalton protein complex involved in binding and entry of donor DNA. J Bacteriol 164:201–206.

74. Hahn J, Albano M, Dubnau D. 1987. Isolation and characterization of Tn917lac-generated competence mutants of Bacillus subtilis. J Bacteriol 169:3104–3109.

75. Albano M, Breitling R, Dubnau DA. 1989. Nucleotide sequence and genetic organization of the Bacillus subtilis comG operon. J Bacteriol 171:5386–5404.

76. Dubnau D. 1991. Genetic competence in Bacillus subtilis. Microbiol Rev. American Society for Microbiology (ASM) https://doi.org/10.1128/mmbr.55.3.395-424.1991.

77. Chung YS, Dubnau D. 1998. All seven comG open reading frames are required for DNA binding during transformation of competent Bacillus subtilis. J Bacteriol 180:41–45.

78. Chung YS, Dubnau D. 1995. ComC is required for the processing and translocation of ComGC, a pilin-like competence protein of Bacillus subtilis. Mol Microbiol 15:543–551.

79. Livingstone CD, Barton GJ. 1993. Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation. Comput Appl Biosci 9:745–756.

80. Chen I, Provvedi R, Dubnau D. 2006. A macromolecular complex formed by a pilin-like protein in competent Bacillus subtilis. Journal of Biological Chemistry 281:21720–21727.

81. Inamine GS, Dubnau D. 1995. ComEA, a Bacillus subtilis integral membrane protein required for genetic transformation, is needed for both DNA binding and transport. J Bacteriol 177:3045–3051.

82. Provvedi R, Dubnau D. 1999. ComEA is a DNA receptor for transformation of competent Bacillus subtilis. Mol Microbiol 31:271–280.

83. Craig L, Forest KT, Maier B. 2019. Type IV pili: dynamics, biophysics and functional consequences. Nat Rev Microbiol. Nature Publishing Group https://doi.org/10.1038/s41579-019-0195-4.

84. Craig L, Li J. 2008. Type IV pili: paradoxes in form and function. Curr Opin Struct Biol. NIH Public Access https://doi.org/10.1016/j.sbi.2007.12.009.

85. Ellison CK, Kan J, Chlebek JL, Hummels KR, Panis G, Viollier PH, Biais N, Dalia AB, Brun Y V. 2019. A bifunctional ATPase drives tad pilus extension and retraction. Sci Adv 5.

86. Campos M, Cisneros DA, Nivaskumar M, Francetic O. 2013. The type II secretion system - a dynamic fiber assembly nanomachine. Res Microbiol 164:545–555.

87. Chen I, Christie PJ, Dubnau D. 2005. The Ins and Outs of DNA Transfer in Bacteria. Science 310:1456.

88. Tønjum T, Freitag NE, Namork E, Koomey M. 1995. Identification and characterization of pilG, a highly conserved pilus-assembly gene in pathogenic Neisseria. Mol Microbiol 16:451–464.

89. Graupner S, Frey V, Hashemi R, Lorenz MG, Brandes G, Wackernagel W. 2000. Type IV pilus genes pilA and pilC of Pseudomonas stutzeri are required for natural genetic transformation, and pilA can be replaced by corresponding genes from nontransformable species. J Bacteriol 182:2184–2190.

90. Aas FE, Wolfgang M, Frye S, Dunham S, Løvold C, Koomey M. 2002. Competence for natural transformation in Neisseria gonorrhoeae: components of DNA binding and uptake linked to type IV pilus expression. Mol Microbiol 46:749–760.

91. Van Schaik EJ, Giltner CL, Audette GF, Keizer DW, Bautista DL, Slupsky CM, Sykes BD, Irvin RT. 2005. DNA binding: A novel function of Pseudomonas aeruginosa type IV pili. J Bacteriol 187:1455–1464.

92. Cehovin A, Simpson PJ, McDowell MA, Brown DR, Noschese R, Pallett M, Brady J, Baldwin GS, Lea SM, Matthews SJ, Pelicic V. 2013. Specific DNA recognition mediated by a type IV pilin. Proc Natl Acad Sci U S A 110:3065–3070.

93. Laurenceau R, Péhau-Arnaudet G, Baconnais S, Gault J, Malosse C, Dujeancourt A, Campo N, Chamot-Rooke J, Le Cam E, Claverys J-P, Fronzes R. 2013. A Type IV Pilus Mediates DNA Binding during Natural Transformation in Streptococcus pneumoniae. PLoS Pathog 9:e1003473.

94. Blair KM, Turner L, Winkelman JT, Berg HC, Kearns DB. 2008. A molecular clutch disables flagella in the Bacillus subtilis biofilm. Science (1979) 320:1636–1638.

95. Ellison CK, Kan J, Dillard RS, Kysela DT, Ducret A, Berne C, Hampton CM, Ke Z, Wright ER, Biais N, Dalia AB, Brun Y V. 2017. Obstruction of pilus retraction stimulates bacterial surface sensing. Science (1979) 358:535–538.

96. Ellison CK, Dalia TN, Vidal Ceballos A, Wang JCY, Biais N, Brun Y V., Dalia AB. 2018. Retraction of DNA-bound type IV competence pili initiates DNA uptake during natural transformation in Vibrio cholerae. Nat Microbiol https://doi.org/10.1038/s41564-018-0174-y.

97. Lam T, Ellison CK, Eddington DT, Brun Y V., Dalia AB, Morrison DA. 2021. Competence pili in Streptococcus pneumoniae are highly dynamic structures that retract to promote DNA uptake. Mol Microbiol 116:381–396.

98. Dubnau D, Blokesch M. 2019. Mechanisms of DNA Uptake by Naturally Competent Bacteria. https://doi.org/101146/annurev-genet-112618-043641 53:217–237.

99. Demchick P, Koch AL. 1996. The permeability of the wall fabric of Escherichia coli and Bacillus subtilis. J Bacteriol 178:768–773.

100. Swoboda JG, Campbell J, Meredith TC, Walker S. 2010. Wall teichoic acid function, biosynthesis, and inhibition. ChemBioChem https://doi.org/10.1002/cbic.200900557.

101. Meima R, Eschevins C, Fillinger S, Bolhuis A, Hamoen LW, Dorenbos R, Quax WJ, Maarten Van Dijl J, Provvedi R, Chen I, Dubnau D, Bron S. 2002. The bdbDC operon of Bacillus subtilis encodes thiol-disulfide oxidoreductases required for competence development. J Biol Chem 277:6994–7001.

102. Pelicic V. 2008. Type IV pili: e pluribus unum? Mol Microbiol https://doi.org/10.1111/j.1365-2958.2008.06197.x.

103. Korotkov K V., Sandkvist M, Hol WGJ. 2012. The type II secretion system: biogenesis, molecular architecture and mechanism. Nature Reviews Microbiology 2012 10:5 10:336–351.

104. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. Nature 2021 596:7873 596:583–589.

105. Chung YS, Breidt F, Dubnau D. 1998. Cell surface localization and processing of the ComG proteins, required for DNA binding during transformation of Bacillus subtilis. Mol Microbiol 29:905–913.

106. Craig L, Taylor RK, Pique ME, Adair BD, Arvai AS, Singh M, Lloyd SJ, Shin DS, Getzoff ED, Yeager M, Forest KT, Tainer JA. 2003. Type IV pilin structure and assembly: X-ray and EM analyses of Vibrio cholerae toxin-coregulated pilus and Pseudomonas aeruginosa PAK pilin. Mol Cell 11:1139–1150.

107. Craig L, Volkmann N, Arvai AS, Pique ME, Yeager M, Egelman EHH, Tainer JA. 2006. Type IV Pilus Structure by Cryo-Electron Microscopy and Crystallography: Implications for Pilus Assembly and Functions. Mol Cell 23:651–662.

108. Horton RM, Cai Z, Ho SN, Pease LR. 2013. Gene splicing by overlap extension: Tailor-made genes using the polymerase chain reaction. Biotechniques 54:528–535.

109. Koo BM, Kritikos G, Farelli JD, Todor H, Tong K, Kimsey H, Wapinski I, Galardini M, Cabal A, Peters JM, Hachmann AB, Rudner DZ, Allen KN, Typas A, Gross CA. 2017. Construction and Analysis of Two Genome-Scale Deletion Libraries for Bacillus subtilis. Cell Syst 4:291-305.e7.

110. Ng D, Harn T, Altindal T, Kolappan S, Marles JM, Lala R, Spielman I, Gao Y, Hauke CA, Kovacikova G, Verjee Z, Taylor RK, Biais N, Craig L. 2016. The Vibrio cholerae Minor Pilin TcpB Initiates Assembly and Retraction of the Toxin-Coregulated Pilus. PLoS Pathog 12:e1006109.

111. Vignon G, Köhler R, Larquet E, Giroux S, Prévost MC, Roux P, Pugsley AP. 2003. Type IV-like pili formed by the type II secreton: Specificity, composition, bundling, polar localization, and surface presentation of peptides. J Bacteriol 185:3416–3428.

112. Durand É, Michel G, Voulhoux R, Kürner J, Bernadac A, Filloux A. 2005. XcpX controls biogenesis of the Pseudomonas aeruginosa XcpT-containing pseudopilus. Journal of Biological Chemistry 280:31378–31389.

113. Cisneros DA, Bond PJ, Pugsley AP, Campos M, Francetic O. 2012. Minor pseudopilin self-assembly primes type II secretion pseudopilus elongation. EMBO J 31:1041–1053.

114. Nguyen Y, Sugiman-Marangos S, Harvey H, Bell SD, Charlton CL, Junop MS, Burrows LL. 2015. Pseudomonas aeruginosa minor pilins prime type IVa pilus assembly and promote surface display of the PilY1 adhesin. Journal of Biological Chemistry 290:601–611.

115. Harwood, C R and Cutting SM. 1990. Molecular Biological methods for BacillusMolecular Biological Methods for Bacillus.

116. Auchtung JM, Aleksanyan N, Bulku A, Berkmen MB. 2016. Biology of ICEBs1, an integrative and conjugative element in Bacillus subtilis. Plasmid 86:14–25.

117. Pedreira T, Elfmann C, Stülke J. 2022. The current state of SubtiWiki, the database for the model organism Bacillus subtilis. Nucleic Acids Res 50:D875–D882.

118. Sun D, Ostermaier MK, Heydenreich FM, Mayer D, Jaussi R, Standfuss J, Veprintsev DB. 2013. AAscan, PCRdesign and MutantChecker: A Suite of Programs for Primer Design and Sequence Analysis for High-Throughput Scanning Mutagenesis. PLoS One 8:e78878.

119. Finkel SE, Kolter R. 2001. DNA as a nutrient: novel role for bacterial competence gene homologs. J Bacteriol 183:6288–6293.

120. Redfield RJ. 1993. Evolution of natural transformation: testing the DNA repair hypothesis in Bacillus subtilis and Haemophilus influenzae. Genetics 133:755–761.

121. Guo M, Wang H, Xie N, Xie Z. 2015. Positive Effect of Carbon Sources on Natural Transformation in Escherichia coli: Role of Low-Level Cyclic AMP (cAMP)-cAMP Receptor Protein in the Derepression of rpoS. J Bacteriol 197:3317.

122. Tanooka H, Takahashi A. 1977. Expression of an excision repair gene in transformation of Bacillus subtilis. Mol Gen Genet 153:129–133.

123. Hoelzer MA, Michod RE. 1991. DNA repair and the evolution of transformation in Bacillus subtilis. III. Sex with damaged DNA. Genetics 128:215–223.

124. Yasbin RE. 1977. DNA repair in Bacillus subtilis. II. Activation of the inducible system in competent bacteria. Mol Gen Genet 153:219–225.

125. Yasbin RE, Fernwalt JD, Fields PI. 1979. DNA repair in Bacillus subtilis: excision repair capacity of competent cells. J Bacteriol 137:391–396.

126. Snitkin ES, Zelazny AM, Montero CI, Stock F, Mijares L, Mullikin J, Blakesley R, Young A, Chu G, Ramsahoye C, Lovett S, Han J, Legaspi R, Sison C, Gregory M, Montemayor C, Gestole M, Hargrove A, Johnson T, Myrick J, Riebow N, Schmidt B, Novotny B, Gupta J, Benjamin B, Brooks S, Coleman H, Ho SL, Schandler K, Smith L, Stantripop M, Maduro Q, Bouffard G, Dekhtyar M, Guan X, Masiello C, Maskeri B, McDowell J, Park M, Thomas P, Murray PR, Segre JA. 2011. Genome-wide recombination drives diversification of epidemic strains of Acinetobacter baumannii. Proc Natl Acad Sci U S A 108:13758–13763.

127. Matthey N, Stutzmann S, Stoudmann C, Guex N, Iseli C, Blokesch M. 2019. Neighbor predation linked to natural competence fosters the transfer of large genomic regions in vibrio cholerae. Elife 8.

128. Winter M, Buckling A, Harms K, Johnsen PJ, Vos M. 2021. Antimicrobial resistance acquisition via natural transformation: context is everything. Curr Opin Microbiol 64:133–138.

129. Blokesch M. 2017. In and out—contribution of natural transformation to the shuffling of large genomic regions. Curr Opin Microbiol 38:22–29.

130. D D. 1991. Genetic competence in Bacillus subtilis. Microbiol Rev 55:395–424.

131. Pelicic V. 2023. Mechanism of assembly of type 4 filaments: everything you always wanted to know (but were afraid to ask). Microbiology (Reading) 169:001311.

132. Ahmed I, Hahn J, Henrickson A, Khaja FT, Demeler B, Dubnau D, Neiditch MB. 2022. Structure-function studies reveal ComEA contains an oligomerization domain essential for transformation in gram-positive bacteria. Nat Commun 13.

133. Hahn J, Desantis M, Dubnau D. 2021. Mechanisms of Transforming DNA Uptake to the Periplasm of Bacillus subtilis. mBio 12.

134. Bergé M, Moscoso M, Prudhomme M, Martin B, Claverys JP. 2002. Uptake of transforming DNA in Gram-positive bacteria: a view from Streptococcus pneumoniae. Mol Microbiol 45:411–421.

135. Draskovic I, Dubnau D. 2005. Biogenesis of a putative channel protein, ComEC, required for DNA uptake: membrane topology, oligomerization and formation of disulphide bonds. Mol Microbiol 55:881–896.

136. Baker JA, Simkovic F, Taylor HMC, Rigden DJ. 2016. Potential DNA binding and nuclease functions of ComEC domains characterized in silico. Proteins 84:1431.

137. Diallo A, Foster HR, Gromek KA, Perry TN, Dujeancourt A, Krasteva P V., Gubellini F, Falbel TG, Burton BM, Fronzes R. 2017. Bacterial transformation: ComFA is a DNA-dependent ATPase that forms complexes with ComFC and DprA. Mol Microbiol 105:741–754.

138. Chilton SS, Falbel TG, Hromada S, Burton BM. 2017. A Conserved Metal Binding Motif in the Bacillus subtilis Competence Protein ComFA Enhances Transformation. J Bacteriol 199.

139. Takeno M, Taguchi H, Akamatsu T. 2011. Role of ComFA in controlling the DNA uptake rate during transformation of competent Bacillus subtilis. J Biosci Bioeng 111:618–623.

140. Yadav T, Carrasco B, Myers AR, George NP, Keck JL, Alonso JC. 2012. Genetic recombination in Bacillus subtilis: a division of labor between two single-strand DNA-binding proteins. Nucleic Acids Res 40:5546.

141. Yadav T, Carrasco B, Hejna J, Suzuki Y, Takeyasu K, Alonso JC. 2013. Bacillus subtilis DprA Recruits RecA onto Single-stranded DNA and Mediates Annealing of Complementary Strands Coated by SsbB and SsbA. J Biol Chem 288:22437.

142. Briley K, Dorsey-Oresto A, Prepiak P, Dias MJ, Mann JM, Dubnau D. 2011. The secretion ATPase ComGA is required for the binding and transport of transforming DNA. Mol Microbiol 81:818.

143. Hahn J, Maier B, Haijema BJ, Sheetz M, Dubnau D. 2005. Transformation Proteins and DNA Uptake Localize to the Cell Poles in Bacillus subtilis. Cell 122:59.

144. Kramer N, Hahn J, Dubnau D. 2007. Multiple interactions among the competence proteins of Bacillus subtilis. Mol Microbiol 65:454–464.

145. Popp PF, Dotzler M, Radeck J, Bartels J, Mascher T. 2017. The Bacillus BioBrick Box 2.0: expanding the genetic toolbox for the standardized work with Bacillus subtilis. Scientific Reports 2017 7:1 7:1–13.

146. Salleh MZ, Karuppiah V, Snee M, Thistlethwaite A, Levy CW, Knight D, Derrick JP. 2019. Structure and properties of a natural competence-associated pilin suggest a unique pilus Tip-associated DNA receptor. mBio 10.

147. Chlebek JL, Denise R, Craig L, Dalia AB. 2021. Motor-independent retraction of type IV pili is governed by an inherent property of the pilus filament. Proc Natl Acad Sci U S A 118.

148. Hsu YP, Hall E, Booher G, Murphy B, Radkov AD, Yablonowski J, Mulcahey C, Alvarez L, Cava F, Brun Y V., Kuru E, VanNieuwenhze MS. 2019. Fluorogenic D-amino acids enable real-time monitoring of peptidoglycan biosynthesis and high-throughput transpeptidation assays. Nat Chem 11:335–341.

149. Briley K, Prepiak P, Dias MJ, Hahn J, Dubnau D. 2011. Maf acts downstream of ComGA to arrest cell division in competent cells of B. subtilis. Mol Microbiol 81:23.

150. Mirouze N, Ferret C, Yao Z, Chastanet A, Carballido-López R. 2015. MreB-Dependent Inhibition of Cell Elongation during the Escape from Competence in Bacillus subtilis. PLoS Genet 11:e1005299.

151. Seitz P, Blokesch M. 2013. DNA-uptake machinery of naturally competent Vibrio cholerae. Proc Natl Acad Sci U S A 110:17987–17992.

152. Kaufenstein M, Van Der Laan M, Graumann PL. 2011. The Three-Layered DNA Uptake Machinery at the Cell Pole in Competent Bacillus subtilis Cells Is a Stable Complex. J Bacteriol 193:1633.

153. Polasek-Sedlackova H, Miller TCR, Krejci J, Rask MB, Lukas J. 2022. Solving the MCM paradox by visualizing the scaffold of CMG helicase at active replisomes. Nature Communications 2022 13:1 13:1–11.

154. Burghard-Schrod M, Kilb A, Krämer K, Graumann PL. 2022. Single-Molecule Dynamics of DNA Receptor ComEA, Membrane Permease ComEC, and Taken-Up DNA in Competent Bacillus subtilis Cells. J Bacteriol 204:572–593.

155. Oliveira V, Aschtgen MS, van Erp A, Henriques-Normark B, Muschiol S. 2021. The Role of Minor Pilins in Assembly and Function of the Competence Pilus of Streptococcus pneumoniae. Front Cell Infect Microbiol 11.

156. Harwood CR, Cutting. 1990. Molecular Biological Methods for Bacillus. Wiley, New York.

157. Sambrook J, Russell D. Molecular Cloning: A Laboratory Manual Third Edition.

158. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature Methods 2009 6:5 6:343–345.

159. Gryczan TJ, Contente S, Dubnau D. 1978. Characterization of Staphylococcus aureus plasmids introduced by transformation into Bacillus subtilis. J Bacteriol 134:318–329.

160. Litovchick L. 2018. Immunoblotting: Transfer of proteins from gels to membranes. Cold Spring Harb Protoc 2018:818–824.

161. Prasanna VR, Ayse Sena Mutlu, Meng CW. 2015. Label-Free Biomedical Imaging of Lipids by Stimulated Raman Scattering Microscopy. Curr Protoc Mol Biol 109:30.3.1-30.3.17.

162. Rudner DZ, Losick R. 2002. A sporulation membrane protein tethers the pro-sigmaK processing enzyme to its inhibitor and dictates its subcellular localization. Genes Dev 16:1007–1018.

163. Burton BM, Marquis KA, Sullivan NL, Rapoport TA, Rudner DZ. 2007. The ATPase SpoIIIE Transports DNA across Fused Septal Membranes during Sporulation in Bacillus subtilis. Cell 131.

164. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji - an Open Source platform for biological image analysis. Nat Methods 9:676–682.

165. Ducret A, Quardokus EM, Brun Y V. 2016. MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. Nat Microbiol 1.

166. Earl AM, Losick R, Kolter R. 2008. Ecology and genomics of Bacillus subtilis. Trends Microbiol 16:269.

167. Lutz C, Erken M, Noorian P, Sun S, McDougald D. 2013. Environmental reservoirs and mechanisms of persistence of Vibrio cholerae. Front Microbiol 4.

168. Green AE, Pottenger S, Monshi MS, Barton TE, Phelan M, Neill DR. 2023. Airway metabolic profiling during Streptococcus pneumoniae infection identifies branched chain amino acids as signatures of upper airway colonisation. PLoS Pathog 19:e1011630.

169. Li J, Egelman EH, Craig L. 2012. Structure of the Vibrio cholerae Type IVb Pilus and stability comparison with the Neisseria gonorrhoeae type IVa pilus. J Mol Biol 418:47–64.

170. Dos Santos Morais R, El-Kirat-Chatel S, Burgain J, Simard B, Barrau S, Paris C, Borges F, Gaiani C. 2020. A Fast, Efficient and Easy to Implement Method to Purify Bacterial Pili From Lacticaseibacillus rhamnosus GG Based on Multimodal Chromatography. Front Microbiol 11:3225.

171. Nogales E, Scheres SHW. 2015. Molecular Cell Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity https://doi.org/10.1016/j.molcel.2015.02.019.

172. Weissenberger G, Henderikx RJM, Peters PJ. 2021. Understanding the invisible hands of sample preparation for cryo-EM. Nature Methods 2021 18:5 18:463–471.

173. Yip KM, Fischer N, Paknia E, Chari A, Stark H. 2020. Atomic-resolution protein structure determination by cryo-EM. Nature 2020 587:7832 587:157–161.

174. Fernandez-Leiro R, Bhairosing-Kok D, Kunetsky V, Laffeber C, Winterwerp HH, Groothuizen F, Fish A, Lebbink JHG, Friedhoff P, Sixma TK, Lamers MH. 2021. The selection process of licensing a DNA mismatch for repair. Nature Structural & Molecular Biology 2021 28:4 28:373–381.

175. Muschiol S, Erlendsson S, Aschtgen MS, Oliveira V, Schmieder P, De Lichtenberg C, Teilum K, Boesen T, Akbey U, Henriques-Normark B. 2017. Structure of the competence pilus major pilin ComGC in Streptococcus pneumoniae. J Biol Chem 292:14134.

176. Lobstein J, Emrich CA, Jeans C, Faulkner M, Riggs P, Berkmen M. 2012. SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. Microb Cell Fact 11:1–16.

177. Froger A, Hall JE. 2007. Transformation of Plasmid DNA into E. coli Using the Heat Shock Method. J Vis Exp 6.

178. Sagen A. 2023. Standard M9 minimal medium.

179. Youngman PJ, Perkins JB, Losick R. 1983. Genetic transposition and insertional mutagenesis in Bacillus subtilis with Streptococcus faecalis transposon Tn917. Proc Natl Acad Sci U S A 80:2305–2309.

180. Grant SGN, Jessee J, Bloom FR, Hanahan D. 1990. Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. Proc Natl Acad Sci U S A 87:4645–4649.

181. Youngman PJ, Perkins JB, Losick R. 1983. Genetic transposition and insertional mutagenesis in Bacillus subtilis with Streptococcus faecalis transposon Tn917. Proc Natl Acad Sci U S A 80:2305–2309.

182. Chilton SS, Falbel TG, Hromada S, Burton BM. 2017. A conserved metal binding motif in the Bacillus subtilis competence protein ComFA enhances transformation. J Bacteriol 199:e00272-017.

183. Lemon KP, Grossman AD. 1998. Localization of bacterial DNA polymerase: evidence for a factory model of replication. Science 282:1516–1519.

184. Studier FW, Moffatt BA. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol 189:113–130.

185. Comella BS, Grossman AD, Comella N, Grossman AD. 2005. Conservation of genes and processes controlled by the quorum response in bacteria: characterization of genes controlled by the quorum-sensing transcription factor ComA in Bacillus subtilis. Mol Microbiol 57:1159–1174.

186. Hilgarth RS, Lanigan TM. 2020. Optimization of overlap extension PCR for efficient transgene construction. MethodsX 7:100759.

187. Kraus A, Hueck C, Gartner D, Hillen W. 1994. Catabolite repression of the Bacillus subtilis xyl operon involves a cis element functional in the context of an unrelated sequence, and glucose exerts additional xylR-dependent repression. J Bacteriol 176:1738.

188. Zuke JD, Erickson R, Hummels KR, Burton BM. 2023. Visualizing dynamic competence pili and DNA capture throughout the long axis of Bacillus subtilis . J Bacteriol https://doi.org/10.1128/JB.00156-23/SUPPL_FILE/JB.00156-23-S0003.AVI.