Regulation and Substrate Selectivity of Peptide Exporters and Their Functional Significance in *Streptococcus pneumoniae*

by

Charles Y. Wang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Microbiology and Immunology)
in the University of Michigan
2019

Doctoral Committee:

Associate Professor Suzanne R. Dawid, Chair
Associate Professor Eric C. Martens
Professor Mary X. O'Riordan
Associate Professor Lyle A. Simmons

Charles Y. Wang

wangchar@umich.edu

ORCID iD: 0000-0003-2060-2378

DEDICATION

To my parents, Xin Li and Chengfei Wang.

ACKNOWLEDGMENTS

First, I'd like to thank my mentor, Suzy Dawid, for being incredibly supportive these past four years. I've learned many invaluable lessons about science from her and appreciate how she pushed me to think about my project in new and creative ways. I am also grateful for all the times she trusted me enough to let me pursue my crazy or long-shot ideas – a lot of which didn't pan out but that only made the ones that did so much more satisfying.

I want to acknowledge my committee, Eric Martens, Mary O'Riordan, and Lyle Simmons, for all their advice and feedback on my project. I'd also like to thank the MSTP, in particular Ron, Ellen, and Justine, for being enormously helpful and supportive and helping to make the dual-degree process go as smoothly as possible.

I also want to thank the many current and former members of the Dawid, Watson, and Weinberg labs with whom it has been a joy to work alongside. Thank you to Winnie Wholey for teaching me everything I needed to know about working with pneumococcus. I must also mention Jennifer Medlin, who has been a terrific lab manager and a great person to talk with about science, TV, books, movies, and just about any other random topic that pops up. I've also had the pleasure of working with four talented undergrads, Nisha Patel, Schyler Bennett, Don Nguyen, and Miguel Disbennett, who have all made invaluable contributions to the work in this thesis.

I need to give a special thank you to my family. I would not be where I am without a lifetime of love and support from my mom and dad. And thank you to my brother, Kevin, with whom I can always talk and laugh about anything.

Lastly, this thesis work was supported by grants from the National Institutes of Health (R01AI101285, T32GM007863, and T32AI007528).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

ABC transporter	ATP-binding cassette transporter
ADP	adenosine diphosphate
ANOVA	analysis of variance
ATP	adenosine triphosphate
BIR	bacteriocin/immunity region
cam	chloramphenicol
CSP	competence-stimulating peptide
dsDNA	double-stranded DNA
DTT	dithiothreitol
gen	gentamicin
GG peptide	double-glycine motif peptide
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HGT	horizontal gene transfer
Luc	firefly luciferase, wildtype from *Photinus pyralis*
MES	2-(*N*-morpholino)ethanesulfonic acid
NBD	nucleotide-binding domain
Nluc	NanoLuc luciferase
OD	optical density
PBS	phosphate-buffered saline
PCAT	peptidase-containing ABC transporter
PCR	polymerase chain reaction
PCV	pneumococcal conjugate vaccine
PEP	peptidase domain
RFluc	firefly luciferase, red-emitting mutant from *Luciola italica*
RLU	relative light units
SD	standard deviation
SDS	sodium dodecyl sulfate
SE	standard error
SHP	short hydrophobic peptide
spc	spectinomycin
SS	signal sequence
ssDNA	single-stranded DNA
str	streptomycin
TBS	Tris-buffered saline
TMD	transmembrane domain

ABSTRACT

Peptidase-containing ABC transporters (PCATs) are a widely distributed family of transporters which export peptide substrates containing a double-glycine motif (GG peptides). In gram-positive bacteria, PCATs secrete pheromones used for cell-to-cell signaling and antimicrobial peptides called bacteriocins used for interbacterial killing and competition. PCATs recognize GG peptides through their N-terminal signal sequences, but little is known about how PCATs distinguish between different GG peptides. The opportunistic pathogen *Streptococcus pneumoniae* (pneumococcus) encodes multiple PCATs. Two of these, ComAB and BlpAB, secrete the quorum-sensing pheromones CSP (competence-stimulating peptide) and BlpC, respectively. CSP induces genetic competence, allowing pneumococcus to incorporate extracellular DNA into its genome via homologous recombination to facilitate DNA repair and horizontal gene transfer. BlpC induces the production of the major family of pneumococcal bacteriocins, the Blp bacteriocins (pneumocins).

Recently, it was reported that ComAB could cross-secrete the BlpC pheromone, mediating crosstalk from the competence regulatory system (*com*) to the pneumocin regulatory system (*blp*). Here, I extend that work to show that BlpAB can also cross-secrete CSP, enabling crosstalk in the *blp* to *com* direction. Moreover, the ability of ComAB and BlpAB to share substrates extends to the pneumocins. While nearly all strains produce functional ComAB and encode pneumocins, only 25% produce functional BlpAB. Cross-secretion of CSP and BlpC by ComAB/BlpAB results in complex patterns of *com* and *blp* regulation which differ between BlpAB$^+$ and BlpAB$^-$ strains. First, BlpAB$^+$ strains can activate competence at lower cell densities and under a greater range of conditions than BlpAB$^-$ strains. Second, BlpAB$^+$ strains can secrete pneumocins independently of competence while BlpAB$^-$ strains can only secrete pneumocins during brief periods of competence activation. Moreover, differences in timing and duration of transporter expression between ComAB and BlpAB allow BlpAB$^+$

strains to secrete greater amounts of pneumocins than BlpAB$^-$ strains. This leads to a pneumocin-mediated competitive advantage for BlpAB$^+$ strains over BlpAB$^-$ strains during nasopharyngeal colonization in mice. Therefore, BlpAB$^+$ strains are aggressors which use pneumocins to kill competitors under a wide range of conditions while BlpAB$^-$ strains are opportunists which primarily use pneumocins to support competence.

The cross-secretion between *com*- and *blp*- regulated peptides led me to examine the role of a previously uncharacterized pneumococcal PCAT, RtgAB, in peptide secretion. RtgAB is encoded by the *rtg* locus next to several GG peptides of unknown function. I determined that *rtg* is regulated by the RtgR/RtgS system, in which RtgS, a SHP (small hydrophobic peptide)-like pheromone with a distinctive Trp-X-Trp motif, is exported then reimported back into the cell to induce *rtg* through the Rgg-family transcription regulator RtgR. An active RtgR/S system provides a competitive fitness advantage in a mouse model of nasopharyngeal colonization. Since ComAB and BlpAB share substrates, I investigated the ability of RtgAB to do the same and found that RtgAB and ComAB/BlpAB preferentially secrete different sets of GG peptides. This selectivity is determined by the GG peptides' signal sequences; ComAB/BlpAB prefers substrates with certain hydrophobic residues at conserved signal sequence positions while RtgAB prefers substrates with a unique motif at the N-terminal end of the signal sequence. These findings illuminate a relatively understudied part of PCAT biology and will help guide future efforts to predict PCAT-substrate pairings. Ultimately, studying PCAT regulation and how they secrete GG peptides will advance our understanding of the many microbial processes dependent on these transporters.

Chapter I: Introduction

1.1 Peptidase-Containing ABC Transporters (PCATs)

Export of polypeptides from their site of synthesis in the cytoplasm to the extracellular space is a fundamental physiological function of all cells. The secretome, the collection of all non-membrane associated proteins secreted from the cell, may comprise up to 20% of an organism's total proteome (1). Bacteria have evolved many different strategies for exporting proteins and peptides (2). One such strategy is the secretion of peptides using a family of ATP-binding cassette (ABC) transporters called peptidase-containing ABC transporters (PCATs).

Like all ABC transporters (3), PCATs are integral membrane proteins that directly couple translocation of substrate across membranes with ATP hydrolysis. Reflective of their specialized roles in peptide export, PCATs contain cytoplasmically localized N-terminal peptidase domains (PEPs) (4), from which they derive their name (Fig. 1.1). These peptidase domains belong to the C39 family of cysteine proteases (4) that are structurally similar to papain-like cysteine proteases (5-7). The peptidase domains are responsible for recognition and proteolytic processing of substrates prior to transport (5-8). In addition to the peptidase domain, each functional PCAT unit contains two transmembrane domains (TMDs) and two cytoplasmically localized nucleotide-binding domains (NBDs) (Fig. 1.1). The TMDs together form the channel through which substrates are translocated across the membrane. The NBDs catalyze ATP hydrolysis, which provides the energy for transport. These domains are usually organized such that a single polypeptide chain contains one of each PEP, TMD and NBD. Two polypeptide chains associate to fulfill the requirement of two TMDs and two NBDs, forming a functional dimer.

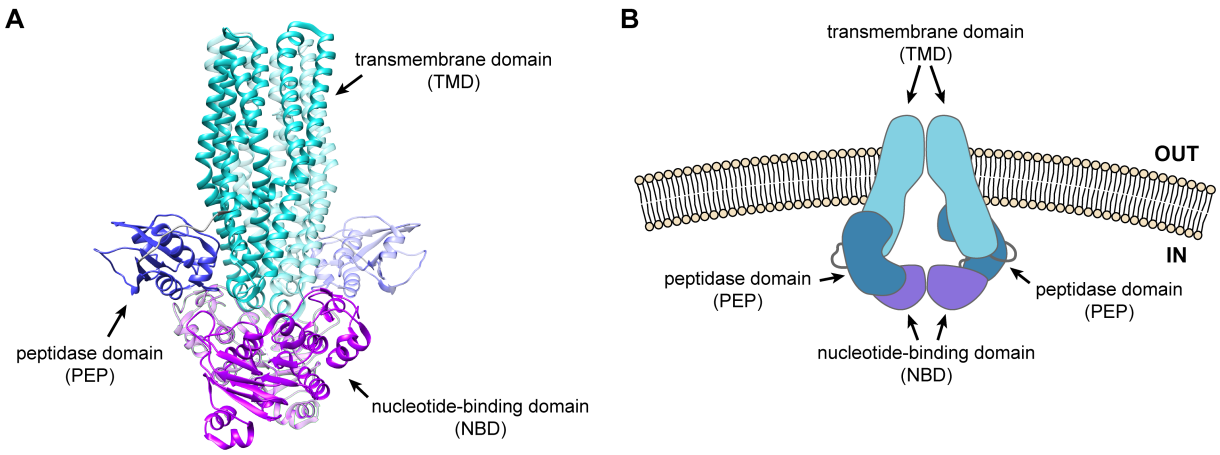

Figure 1.1. Structure of Peptidase-containing ABC Transporters (PCATs). (A) Crystal structure of PCAT1 from *Clostridium thermocellum* in its ATP-free state (6). The full structure is a homodimer of two polypeptide chains; one chain is colored solid and the second is semi-transparent. The chains are colored by domain. **(B)** Diagrammatic representation of the domain structure and topology of PCATs with respect to the cell membrane.

1.1.1 Functional Roles

PCATs are found in both gram-negative and gram-positive bacteria. In gram-negative bacteria, they are part of type 1 secretion systems that secrete antimicrobial peptides called bacteriocins (9). These systems transport their substrates across the inner and outer membranes directly from the cytoplasm to the extracellular space, bypassing the periplasm entirely. Some gram-negative PCATs contain peptidase domains that do not proteolytically cleave substrate but nonetheless are required for substrate recognition (9). These transporters are also associated with type 1 secretion systems and secrete larger proteins instead of peptides.

In contrast to their gram-negative counterparts, gram-positive PCATs function either alone or with a single additional accessory protein (10) instead of as part of a larger secretion complex. They play important roles in bacteriocin secretion and quorum sensing. The remainder of this section will focus on PCATs found in gram-positive bacteria.

The most well-known function of PCATs is to assist in the biosynthesis of bacteriocins. Bacteriocins are ribosomally synthesized antimicrobial peptides produced by bacteria to kill or otherwise inhibit the proliferation of other, usually closely related,

bacteria (11). Bacteriocins are a key part of bacterial adaptation and ecology; they are one of many weapons bacteria deploy against competitors when vying for limited resources (12, 13). From a biotechnology standpoint, bacteriocins have found success in the food industry as preservatives (11). There is also ongoing research into using bacteriocins as anti-bacterial therapeutics, either alone or in combination with traditional antibiotics, to combat the growing prevalence of antibiotic resistance (11).

Bacteriocin-secreting PCATs have been found in numerous gram-positive species. Some examples include the transporters that secrete the Blp bacteriocins from *Streptococcus pneumoniae* (14), the Sil bacteriocins in Streptococcus Anginosus group species (15), the mutacins from *Streptococcus mutans* (16), enterocin A (17) and the NKR-5-3 enterocins (18) from *Enterococcus faecium*, avicin A from *Enterococcus avium* (19), sakacin A from *Lactobacillus sake* (20), lacticin 481 (21) and lactococcins G (4) and Q (22) from *Lactococcus lactis*, and nukacin ISK-1 from *Staphylococcus warneri* (23). A subset of these PCATs also secrete the peptide pheromones that activate expression of their associated bacteriocins through quorum-sensing systems. These pheromone-secreting PCATs include BlpA from *S. pneumoniae* (14, 24), SilE from Streptococcus Anginosus group species (15), EnkT from *E. faecium* (18, 25), and AvcT from *E. avium* (19).

In streptococci, PCATs also participate in the regulation of genetic competence. Genetic competence is the ability of bacteria to take up and incorporate exogenous DNA into their genomes. Competence confers many advantages to bacteria; it allows for horizontal gene transfer, increasing adaptability and diversity, and it offers a pathway for homologous DNA repair, increasing tolerance to mutagens and DNA-damaging stress (26). In *S. pneumoniae* and other Streptococcus Mitis group species, the PCAT ComA secretes the pre-pheromone ComC (27). The active form of ComC, CSP (competence-stimulating peptide), serves as the primary activating signal of genetic competence in these bacteria (28, 29). The non-Mitis group streptococcus *S. mutans* does not use CSP to regulate competence. Nonetheless, in *S. mutans* the PCAT CslA still modulates competence activation by secreting the pheromone MIP (mutacin-inducing peptide), which indirectly activates competence through a poorly defined mechanism (16, 30, 31).

Finally, in *S. pneumoniae* PCATs have been implicated in the secretion of signaling peptides linked to virulence and biofilm formation. The orphan PCAT substrate vp1, whose transporter has not been experimentally determined, promotes biofilm formation and invasive disease (32). Another PCAT substrate, BriC, is secreted by ComA and links genetic competence to enhancement of biofilm formation and nasopharyngeal colonization (33).

1.1.2 Substrate Recognition

Isolated PEPs of various PCATs bind substrate in the absence of the remainder of the transporter protein (4, 6-8, 21). Additionally, a truncated form of PCAT1 from *Clostridium thermocellum* lacking PEP failed to bind substrate (6). These data show that in many cases PEP is necessary and sufficient for substrate binding and is likely the primary driver of substrate recognition in PCATs.

PCAT substrates have an N-terminal signal sequence, or leader peptide, that terminates in a conserved double-glycine (GG) motif, with occasional GA or GS variants (34) (Fig. 1.2A). For this reason, these substrates are called double-glycine or GG peptides. Cleavage of the substrate to remove the signal sequence occurs directly after the GG motif. Mutation of the GG motif either abolishes or severely inhibits substrate cleavage (21), explaining its near-perfect conservation in this family of peptides.

In addition to the GG motif, the signal sequences of GG peptides contain several other conserved residues. A common numbering convention defines position −1 as the residue on the N-terminal side of the scissile bond during signal sequence cleavage and position +1 as the residue on the C-terminal side of the scissile bond. Using this convention, positions −4, −7, −12, and −15 tend to be restricted to residues with hydrophobic side chains (34). Of these four positions, −7 and −12 are the most stringently conserved and almost exclusively contain hydrophobic amino acids. Position −4 occasionally contains residues with small polar side chains such as threonine (33) or cysteine (22). Position −15 is the most relaxed; polar (7, 19) or even charged (18, 19, 35) side chains can be found at this position. Consistent with these four residues playing key roles in substrate recognition by PCATs, mutation of any of them in ComC from *S. pneumoniae* substantially decreased binding to and cleavage by ComA PEP compared to the wild-type peptide (8). Similarly, mutating positions −4, −7, and −12 in

substrates of LahT from *Lachnospiraceae bacterium* substantially decreased cleavage efficiency by LahT PEP (7). Studies of other PCAT-substrate pairs have also identified a subset of these four positions to be important for substrate cleavage (21, 36).

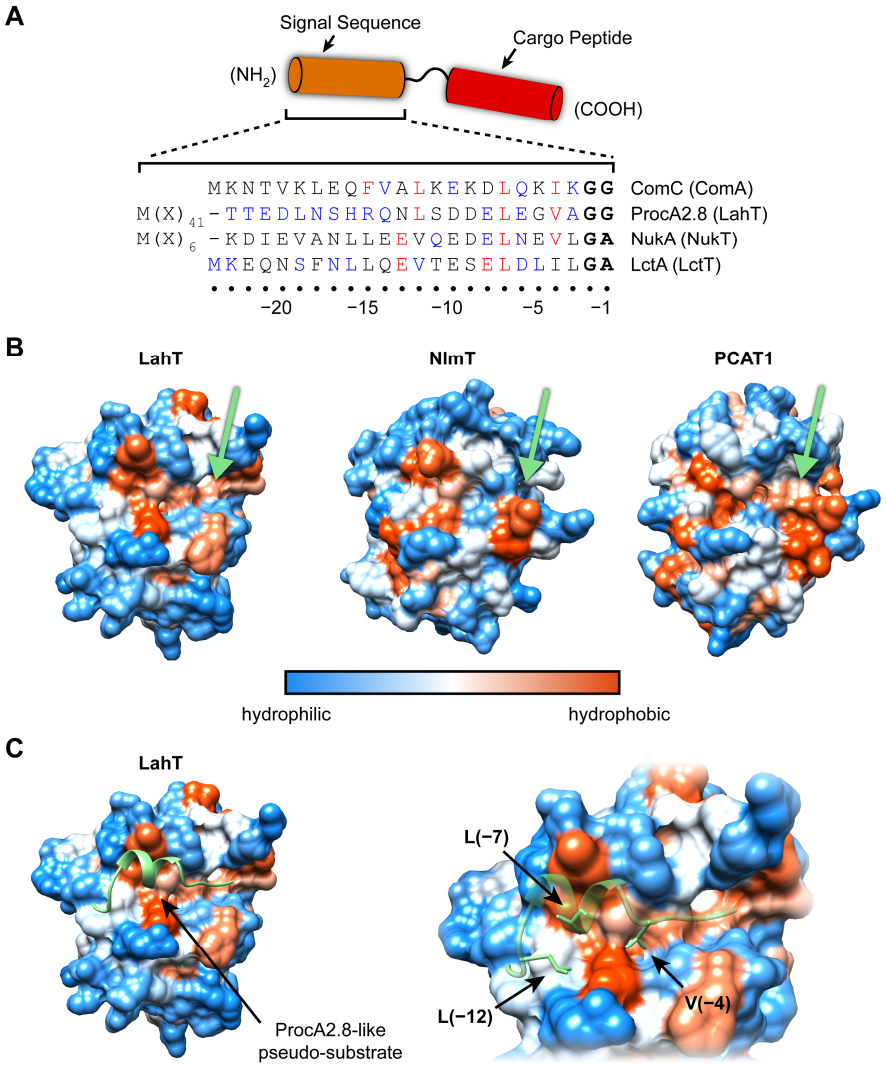

Figure 1.2. PCATs recognize substrates via PEP-signal sequence interactions. (A) Structure of GG peptides. Signal sequences from selected GG peptides are shown. The double-glycine motif is bolded. Residues found to be important for binding to and/or cleavage by the transporter listed in parentheses are highlighted in red. Residues found to be dispensable for binding/cleavage are highlighted in blue. **(B)** Hydrophobicity surface maps of PEP from three different PCATs (5-7). The green arrows point to the active sites. The hydrophobic groove predicted to interact with substrate can be seen in each structure as a red/orange patch to the left of the active site. **(C)** PEP of LahT in complex with a pseudo-substrate (7). Left: Residues −13 to −1 of a ProcA2.8-like pseudo-substrate (green ribbon) bound to PEP of LahT (hydrophobicity surface). Right: Close-up of the binding pocket with pseudo-substrate bound to PEP; labels point to the side chains of the hydrophobic residues in the substrate important for binding to and cleavage by PEP.

Other positions in the signal sequences of GG peptides are relatively non-conserved. The region between positions −7 and −12 tends to be rich in charged and polar residues, but amino acid identity at any particular position varies greatly across different peptides. Mutating substrates at these non-conserved positions did not substantially affect binding or cleavage by ComA or LahT PEP (7, 8). In contrast, some of these positions have been shown to be important for cleavage by NukT and LctT (21, 36).

Crystal structures of PEP from NlmT (5), PCAT1 (6), and LahT (7), the latter in complex with a pseudo-substrate, provide insights into PEP-substrate interactions during substrate binding. Analysis of all three structures reveal that the entrance to the active site of PEP is narrow (Fig. 1.2B). This explains the near-absolute requirement for the GG motif at positions −1 and −2; residues with larger side chains would be sterically blocked from the entrance and prevent the correct positioning of the substrate within the active site. Additionally, all three PEPs have hydrophobic patches on their surfaces near the active site (Fig. 1.2B). This hydrophobic patch could potentially interact with the four conserved hydrophobic residues in the signal sequence of substrate. In silico modeling of the ComA PEP-ComC complex from *S. pneumoniae* yielded probable docking poses in which the four conserved hydrophobic residues in ComC directly interact with the hydrophobic patch in ComA PEP (5). The structure of LahT in complex with a pseudo-substrate also shows residues −4, −7, and −12 of substrate making close contact with the hydrophobic patch of LahT PEP (7) (Fig. 1.2C). In both complexes, the side chains of the non-conserved residues of the substrate signal sequence are generally facing away from PEP and are exposed to solvent.

While the structures of the ComA and LahT complexes are mostly similar, there are notable differences. First, the entrance to the LahT PEP active site is more open than in ComA PEP. This allows the LahT substrate to occupy a more direct path between the hydrophobic face of LahT PEP and the active site. Meanwhile, ComC must bend around an obstacle that protrudes out between the hydrophobic patch of ComA PEP and the active site. The slightly different poses adopted by the two substrates results in different interactions between the conserved substrate residues at positions −4, −7, and −12 and PEP. Each ComC residue interacts with a more proximal (relative

to the active site) portion of the hydrophobic patch of PEP compared to the same residue in the LahT substrate. As a result, residue −15 of ComC can maintain contact with ComA PEP, but the corresponding residue of the LahT substrate cannot do the same with LahT because residue −12 already occupies the most distal portion of the hydrophobic patch. This offers an explanation for why truncation of residues N-terminal to position −14 in the LahT substrate has no appreciable effect on cleavage by LahT PEP and for why LahT PEP can cleave substrates with divergent, non-hydrophobic amino acids such as asparagine, glutamine, and aspartate at position −15 (7). These residues likely remain detached from LahT PEP during binding and do not contribute to substrate recognition. Moreover, the general finding that position −15 is the least conserved position out of the four hydrophobic residues in the signal sequence of GG peptides can be rationalized if peptides with divergent residues at that position are secreted by LahT-like PCATs.

To date, most research on PCAT substrate recognition has focused exclusively on PEP. While PEP is both necessary and sufficient for substrate recognition, it is not known whether the TMD and NBD participate in an accessory role. Decreased rates of substrate cleavage were reported for isolated PCAT1 PEP compared to full-length PCAT1 (6). However, it is unclear if the decrease resulted from decreased substrate binding or decreased catalytic efficiency. Overlaying the structure of the LahT-substrate complex on the full-length PCAT1 structure reveals that the signal sequence of substrate while bound to PEP is in a position to make contacts with both NBDs of the transporter dimer (7). This leaves open the possibility of the NBDs in full-length transporters interacting with substrate, possibly helping to stabilize the substrate-PEP interaction or helping the transporter to adopt a specific conformation to facilitate substrate cleavage and transport.

The contribution of substrate residues located C-terminal of the cleavage site to substrate recognition is also relatively understudied. Mutating positions +1 and +3 of *S. pneumoniae* ComC to alanine from glutamate and arginine, respectively, results in only modest alterations to binding and cleavage by ComA PEP (8). Similarly, mutation of multiple residues C-terminal to the cleavage site in the lacticin 481 precursor peptide did not prevent cleavage by LctT PEP (21). In contrast, PCAT1 PEP failed to bind a

chimeric pseudo-substrate consisting of the signal sequence of the natural PCAT1 substrate fused to an unrelated lanthanide-binding peptide (6). This topic is especially salient for transporters that secrete peptides that undergo post-translational modifications before secretion. An informative case is the PCAT NukT from *S. warneri*. NukT secretes the post-translationally modified lanthionine-containing bacteriocin (lantibiotic) nukacin ISK-1. If nukacin does not receive the appropriate amino acid modifications – all of which occur at residues outside of the signal sequence – it is not cleaved by full-length NukT (23). However, isolated PEP of LctT, a different lantibiotic transporter from *L. lactis*, could cleave the unmodified form of its substrate (21). It is possible that NukT and LctT recognize substrates through different mechanisms. Alternatively, the authors of the NukT study suggest that the discrepancy between NukT and LctT can be explained if selectivity for modified peptides requires the full-length transporter. Unfortunately, this hypothesis could not be directly tested since isolated NukT PEP cannot cleave either modified or unmodified substrate (23, 37). It is clear that substrate residues outside of the signal sequence can affect substrate recognition. However, more work is needed to elucidate the specific details of these extra-signal sequence requirements.

While progress has been made in determining how PCATs recognize and bind to GG peptides, little is known about how PCATs distinguish between different GG peptides. In general, individual PCATs have been shown to tolerate a wide range of GG peptide substrates. Processing or secretion of heterologous GG peptides has been reported for multiple PCATs. LahT PEP can cleave the heterologous substrates ProcA and AzoA (7). LaqD from *L. lactis* str. QU4 can secrete LagA and LagB, GG peptides encoded by a different strain of *L. lactis* (22). EnkT from *E. faecium* can secrete pseudo-substrates with signal sequences derived from GG peptides from several different species (38). In *S. pneumoniae* two different PCATs, ComA and BlpA, can both secrete the same GG peptide, the bacteriocin-regulating pheromone BlpC (39, 40). This last case highlights how understanding PCAT substrate selectivity can lead to important insights into the physiology of naturally occurring bacterial strains. The substrate selectivity (or lack thereof) of ComA and BlpA directly influences how bacteriocin production is regulated. Since only a subset of *S. pneumoniae* strains encode the active

form of BlpA (14) while ComA is nearly universally conserved, bacteriocin regulation differs among different strains based on whether they have one or both transporters.

Since all GG peptides have similar signal sequences with relatively minor deviations at conserved residues, it is perhaps not surprising that PCATs can secrete heterologous substrates. However, this ability is not universal. For example, PCAT1 does not cleave a heterologous substrate from *Clostridium acetobutylicum* or ComC from *S. pneumoniae* (6). In another example, the same study that reported cross-secretion of LagA/B by LaqD also found that the mirrored pairing of LagD (the natural transporter of LagA/B) with LaqA and LaqB (the natural substrates of LaqD) did not yield detectable amounts of secretion (22). The latter example is puzzling due the high degree of conservation between the pairs LagA/LaqA and LagB/LaqB (41); the signal sequences of the former pair are identical while those of the latter pair only differ at 4 of 26 positions, with none of the divergent residues located at conserved positions. When differences in substrate selectivity cannot be explained by differences at the conserved positions in the signal sequence, then one must look to non-conserved positions or even residues outside the signal sequence. One study explored whether LctT PEP discriminated between substrates based on non-conserved signal sequence residues; the authors found that swapping these residues in the native substrate for those found in a heterologous substrate did not affect cleavage by LctT PEP (21). This suggests that at least for this PCAT the non-conserved signal sequence residues are unimportant for substrate selection. Further investigation is needed to determine if this holds true for other PCATs as well.

1.1.3 Mechanism of Transport

The overall mechanism of peptide transport by PCATs involves cleavage of the precursor peptide by PEP to remove the N-terminal signal sequence followed by secretion of the mature peptide through the channel formed by the TMDs, with the energy required for secretion supplied by ATP hydrolysis by the NBDs (4) (Fig. 1.3). Since the establishment of PCATs as a distinct family of ABC transporters, biochemical and structural studies have gradually filled in details of how PCATs carry out and coordinate the individual steps of transport.

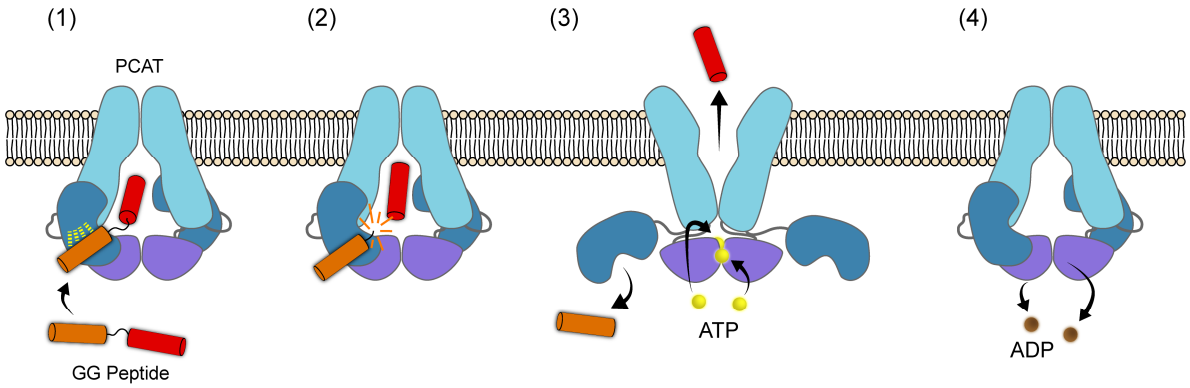

Figure 1.3. PCAT transport mechanism. (1) GG peptide substrate binds to transporter via signal sequence-PEP interactions (yellow dashed lines). (2) PEP cleaves substrate, liberating cargo peptide from the signal sequence. Cargo peptide is loaded into the transporter channel. (3) ATP binding triggers conformational change in transporter, releasing cargo peptide to the extracellular environment. PEP dissociates from the rest of the transporter and cleaved signal sequence is released back into the cytoplasm. (4) NBDs hydrolyze ATP to ADP, which then dissociate from the transporter. This triggers another conformational change to prime the transporter for another round of substrate processing and secretion.

The first step of transport is substrate recognition and binding, which the previous section covers in detail. After binding, PEP cleaves the substrate at the peptide bond directly after the GG motif. The earliest studies of PEP indicated that it was a cysteine protease (4), and the crystal structure of isolated *S. mutans* NlmT PEP confirmed it shared structural similarity to the family of papain-like cysteine proteases (5). Crystal structures of the PEPs from PCAT1 (6) and LahT (7), the former as part of the full-length transporter, showed little structural deviation from each other or from NlmT PEP despite sharing relatively low primary sequence similarity. This indicates that PEP from all PCATs adopt the same general structure and use the same mechanism for substrate cleavage.

PEP employs a catalytic triad of cysteine, histidine, and aspartate in the active site (5, 6). The cysteine and the histidine form a thiolate-imidazolium ion pair, which is stabilized by the aspartate. In the general reaction mechanism for cysteine proteases (42), the deprotonated catalytic cysteine initiates a nucleophilic attack on the carbonyl carbon of the amino acid at position −1 (the second conserved glycine in GG peptides) to form a tetrahedral intermediate. In PEP, the oxyanion of the tetrahedral intermediate is stabilized by either a conserved glutamine in the active site or the catalytic histidine (5, 7). Next, the intermediate spontaneously collapses to release the mature peptide,

leaving the signal peptide attached to the enzyme (acyl-enzyme intermediate) through a thioester bond. The signal peptide is then released from the enzyme through a hydrolysis reaction that also regenerates the enzyme.

While isolated PEP from several PCATs have been shown to independently bind and cleave substrate (5, 7, 8, 21, 28), other studies have described a requirement for full-length transporter for substrate cleavage (23, 37). A study on PCAT1 (6) offers a potential explanation for this discrepancy. While isolated PEP from PCAT1 was able to cleave substrate, the rate of cleavage was five-fold lower than that of full-length, ATP-free PCAT1. Additionally, the ATP-bound form of full-length PCAT1 showed a rate of substrate cleavage comparable to that of isolated PEP. In comparing the structures of ATP-free vs. ATP-bound PCAT1, the authors of the study found substantial differences. In addition to expected conformational changes in the NBDs, where ATP was bound, the structure of ATP-bound PCAT1 did not resolve PEP, indicating that PEP is flexibly detached from the rest of the transporter. This contrasts with the structure of ATP-free PCAT1, in which each PEP of the dimer is locked in a position that puts it close to the interface between the NBDs and TMDs. These data suggest that the close-in orientation of PEP relative to the rest of the transporter promotes efficient substrate cleavage. Since neither isolated PEP nor PEP in ATP-bound full-length transporter is closely associated with the rest of the transporter, they cleave substrate more slowly. If this model is applied to PEPs from other PCATs, we would expect those that can independently cleave substrate are those that can partially tolerate separation from the rest of the transporter (but still cleave substrate more slowly when in the separated state). Meanwhile isolated PEPs that do not cleave substrate are those that cannot tolerate separation from the rest of the transporter.

The next step of transport after substrate cleavage is loading of the mature peptide into the channel formed by the TMDs. The structure of full-length PCAT1 (6) offers insights into how this is achieved. When PCAT1 is in the ATP-free state, each PEP is positioned so that the proteolytic active site sits next to one of two lateral openings in the cytoplasmic end of the TMD channel. Overlaying the LahT-substrate complex over the ATP-free PCAT1 structure reveals that bound substrate is properly positioned to direct the C-terminal end of the substrate into the lateral opening (7).

Therefore, the substrate is primed to release its C-terminal, mature fragment directly into the TMD channel after PEP cleaves off the signal sequence. Moreover, during cleavage the C-terminal portion of the substrate must thread through a narrow channel extending from the PEP active site to the lateral opening in the TMD channel. Since the signal sequence would presumably stay temporarily bound to PEP after cleavage, it may serve as a backstop, occluding the entrance to the narrow channel described above to ensure the mature fragment does not slide backwards out of the TMD channel (7).

The final step of transport is secretion of mature substrate out of the cell. The generally accepted model of transport in ABC transporters is the alternating access model (3). In this model, ATP hydrolysis provides the energy for the transporter channel to alternate between an inward-facing (open to the cytoplasm) and an outward-facing (open to the extracellular space) conformation. Crucially, the channel is never open at both ends at the same time. The ATP-free and ATP-bound structures of PCAT1 are consistent with the alternating access model (6). The ATP-free structure represents the inward-facing conformation. Curiously, the ATP-bound structure has a completely closed TMD channel, open at neither end. Since the structure was obtained in the absence of substrate, the authors of the study speculate that interactions with the substrate in the channel would trigger the full transition to the outward-facing conformation in the ATP-bound state. After the substrate has been secreted, ATP hydrolysis and dissociation of ADP from the NBDs would revert the transporter to the inward-facing state, ready to load another substrate into the channel.

The alternating access model necessitates that the entire substrate must fit into the TMD channel during transport. A logical consequence of this is that any given PCAT can only accommodate mature substrates up to a given length/size. Besides this obvious limitation, little is known about what other features of mature substrates affect loading into the TMD channel and secretion out of the cell after signal sequence cleavage. For example, the questions of whether channel loading is restricted by general properties of the mature substrate other than size such as charge and hydrophobicity, and whether specific interactions between the mature substrate and the channel-lining residues of the transporter play a role remain unanswered.

One of the important questions regarding PCAT function is how or if substrate cleavage is coordinated with secretion of the mature fragment. The deduced transport mechanism of PCAT1 (6) offers a partial answer. Substrate cleavage only occurs efficiently when PCAT1 is in the ATP-free state. In this state, the TMD channel is open to the cytoplasm through two lateral openings and PEPs are bound closely to the TMDs to position the substrate close to these openings. Binding of ATP to PCAT1 triggers a conformational change in the entire transporter that closes the cytoplasmic TMD channel openings and dissociates PEP from the rest of the transporter. In this state PEP cannot efficiently cleave substrate. Therefore, substrate cleavage by PEP preferentially occurs when: 1) the TMD channel is open to the cytoplasm and ready to accept loading of substrate, and 2) PEP is correctly oriented to pass the mature substrate fragment through the TMD channel openings after cleavage. This provides a plausible explanation for how PCATs regulate the activity of PEP to prevent wasted proteolytic cycles when the transporter is not ready to secrete cleaved substrate.

A related issue to the one above is how PCATs regulate the ATPase activity of the NBDs to prevent wasted cycles of ATP hydrolysis when there is no substrate loaded into the TMD channel to secrete. Other ABC transporters solve this problem by coupling ATP hydrolysis with substrate binding. For example, the two NBDs of the maltose transporter dimer remain dissociated from each other and cannot hydrolyze ATP until substrate binding triggers dimerization of the NBDs (43). In contrast, the NBDs of PCAT1 remain in a dimerized, ATP hydrolysis-competent state even in the absence of substrate (6). Consistent with the structural data, biochemical analysis of PCAT1 showed that addition of substrate did not increase ATPase activity (6). Therefore, it is unclear if PCAT1 regulates ATPase activity to prevent unproductive cycling in the absence of substrate. A different result was obtained with another PCAT, the lantibiotic-secreting NukT from *L. lactis* (44). Unlike PCAT1, addition of substrate to NukT increases ATPase activity five-fold. This enhancement of ATPase activity could be partially blocked by inhibition of substrate cleavage, suggesting that either the act of substrate cleavage by PEP per se or loading of cleaved substrate into the TMD channel stimulates ATPase activity in NukT. Since these two studies provide the only currently available data on regulation of ATPase activity by full-length PCATs, further

investigation is needed to determine which model (regulation vs. no regulation) is more common among PCATs and the mechanism underlying regulation in those transporters that use it.

1.1.4 Significance of PCATs

PCATs are a widely distributed class of ABC transporters, found in both gram-negative and gram-positive bacteria. In the latter they primarily function to secrete antimicrobial bacteriocins and their regulatory pheromones. In certain species, including *S. pneumoniae*, PCATs also secrete the pheromones that control activation of genetic competence. Some PCATs also secrete other signaling peptides that promote colonization and virulence. Detailed knowledge of PCATs is required to fully understand these diverse pathways, many of which are found in clinically significant human commensals and pathogens.

Interest in PCATs also extends beyond their native, physiological roles in naturally occurring bacteria. PCAT substrates share a conserved N-terminal signal sequence that is cleaved off during transport. However, the active, mature portions of GG peptides that are secreted out of the cell show great variability. This reflects how PCATs have evolved to transport a diverse range of substrates using the same general mechanism by exploiting the conserved signal sequences as a simple, convenient adapter. The flexibility and modularity of this system makes it an attractive tool to use for synthetic biology (e.g. removal of peptide tags) and the heterologous synthesis of peptide natural products (7, 21). Developing PCATs for these purposes will require a deeper understanding of their biochemistry, especially in the areas of substrate recognition and selectivity.

1.2 *Streptococcus pneumoniae*

Streptococcus pneumoniae (pneumococcus) is an encapsulated gram-positive bacterium that naturally colonizes the human nasopharynx. Pneumococcus colonizes up to 60% of young children (45, 46). As many as half of those who are colonized carry multiple pneumococcal strains (47). Occasionally, pneumococcus escapes from the nasopharynx to invade other sites in the body to cause a wide variety of infections such

as otitis media, sinusitis, pneumonia, meningitis, and bacteremia. *S. pneumoniae* is the most common cause of community-acquired pneumoniae in adults, accounting for at least a quarter of cases (48). The greatest disease burden, however, rests in young children. Pneumococcus-related deaths in children under 5 years of age total more than 800,000 per year worldwide and account for 8% of all deaths in this age group (49). The distribution of pneumococcal disease is skewed heavily toward underdeveloped regions such as Africa, South America, and Central and Southeast Asia (49). Moreover, the true burden of pneumococcal disease is almost certainly larger than what published figures indicate due to incomplete surveillance and reporting and limited sensitivity of diagnostic tests (49).

Since 2000, the pneumococcal conjugate vaccine (PCV) has been available for use in young children (50). PCV directs an immune response against the capsule of pneumococcus. The initial formulation, PCV-7, protected against seven capsule types (serotypes) commonly associated with disease-causing strains. In 2010 in the United States, PCV-7 was replaced by PCV-13, which added six additional serotypes. Both PCV-7 and PCV-13 were effective in reducing the incidence of invasive pneumococcal disease (IPD) in children (51-54). PCV also substantially reduced the colonization rates of vaccine serotypes in children, but overall colonization rates have remained stable due to the expansion of non-vaccine serotypes (55-59).

Prior to the introduction of PCV, antibiotic resistance in *S. pneumoniae* was steadily on the rise (60, 61). While it was expected that PCV would decrease the prevalence of antibiotic-resistant pneumococcus due to elimination of the serotypes with the highest rates of resistance and reduction of antibiotic consumption as a result of decreased disease incidence, actual results have been mixed. Some studies have reported decreasing rates of resistance in the vaccine era (54, 59, 62), while others have found no such effect (58, 63) or mixed trends for different classes of antibiotics (53). In general, PCV has had greater success decreasing antibiotic resistance in IPD than decreasing resistance rates in commensal, colonizing pneumococcal isolates.

It cannot be denied that PCV has been enormously effective at preventing IPD. However, the vaccine has several notable limitations. First, its high cost presents a barrier to implementation in less developed countries, many of which have the highest

burdens of pneumococcal disease (64). Second, it only provides coverage of a limited number of the more than 90 pneumococcal serotypes (65), leaving open the possibility of a resurgence in disease incidence driven by non-vaccine serotypes. In fact, this scenario was realized after the introduction of PCV7. The non-PCV7 serotype 19A rapidly expanded, and increases in IPD incidence caused by 19A partially blunted the vaccine's effectiveness (66). Serotype 19A is now covered by the updated PCV13, but the threat of serotype replacement persists, especially since the reservoir of colonizing pneumococcus strains remains intact due to the ineffectiveness of the vaccine at reducing overall colonization rates. In fact, the incidence of invasive disease caused by serotype 35B strains, which are not covered by PCV13, is currently increasing in the United States; serotype 35B now causes more IPD than any other serotype (67). Given these facts, improving our understanding of how pneumococcal populations thrive in the face of selective pressures to colonize the nasopharynx remains a high priority.

1.3 Genetic Competence in *Streptococcus pneumoniae*

Pneumococcus is a naturally competent organism (68). As such, it can take up and incorporate exogenous DNA into its genome in a process called transformation. This ability was exploited in a series of seminal experiments in the first half of the 20th century to identify DNA as the molecule responsible for carrying genetic information (68, 69).

During transformation (Fig. 1.4, middle), pneumococcus (the recipient) binds extracellular double-stranded DNA (dsDNA) via a type IV pilus (70). The dsDNA (also known as donor DNA) then passes to an import complex where it is converted to single-stranded DNA (ssDNA) and imported into the cell (71). Once internalized, ssDNA is bound by the single-stranded DNA binding protein SsbB, which protects ssDNA from degradation (72). Next, a protein called DprA loads the recombinase RecA onto the ssDNA, allowing RecA to facilitate homologous recombination between the donor ssDNA and the recipient's chromosome (73, 74).

Because integration of donor DNA into the recipient's chromosome relies on homologous recombination, donor and recipient must share a high degree of sequence similarity for efficient transformation to occur. Transformation efficiency decreases

exponentially with increasing sequence divergence; at 15% sequence divergence, roughly the amount seen between *S. pneumoniae* and the most distantly related Mitis group streptococci, transformation efficiency decreases almost 1000-fold compared to using isogenic donor DNA (75). This effectively limits the range of naturally occurring donor DNA compatible for transformation to DNA originating from relatively closely related bacteria. Besides the homology requirement, however, there are no known restrictions on compatible donor DNA. There is no enforcement of specific sequence motifs, such as the DNA uptake sequences required for transformation in *Haemophilus* (76) and *Neisseria* (77), in the donor DNA. Nor are specific methylation patterns required, since conversion of donor DNA to ssDNA prior to internalization, coating of internalized ssDNA by SsbB, and competence-specific methylation of ssDNA (78) all work together to protect the donor DNA from restriction endonucleases.

Pneumococcus can replace large portions of its genome through transformation. Analyses of the genomes of naturally occurring strains have uncovered evidence of multiple recombination events occurring within the same transformation episode, with individual recombination events involving segments of DNA as large as 70 kilobases (79-81). In vitro studies have corroborated the ability of pneumococcus to recombine multiple pieces of DNA per transformation as well as its ability to recombine DNA fragments on the order of 10+ kilobases (82, 83).

While many bacteria can be induced to undergo transformation under artificial laboratory conditions, competence and transformation is part of pneumococcus's natural life cycle. Competence is metabolically expensive (84) and transformation disrupts normal cell growth and proliferation (85). Therefore, pneumococcus has an interest in regulating competence so that it is only active when it is most likely to be of benefit. How pneumococcus controls competence has been and continues to be extensively studied. The next two sections will provide an overview of what is known about this topic.

1.3.1 Regulation of Competence by the com *System*

The default state in pneumococcus is for competence to be turned off. The quorum-sensing system *com* directly regulates competence by controlling the switch from the competence-off to the competence-on state (Fig. 1.4, left). In the first step of

com regulation, cells produce the double-glycine precursor peptide, ComC (28), which is processed and secreted out of the cell by the PCAT ComA in conjunction with an accessory transport protein ComB (27, 86). The mature peptide pheromone, now called CSP (competence-stimulating peptide), accumulates extracellularly. Once CSP reaches a threshold concentration, it signals through the two-component system ComDE (87, 88). CSP binds to the cell-surface receptor and histidine kinase ComD, promoting autophosphorylation and subsequent transphosphorylation of the response regulator ComE (89). Phospho-ComE is induced to dimerize and bind to conserved direct repeats in the promoter regions of target genes to activate their transcription (89-91). The set of genes upregulated by ComE is called the early (competence) genes and includes *comAB* and *comCDE* (92, 93). The fact that CSP upregulates its own expression as well as the proteins responsible for its secretion and sensing results in a positive feedback loop.

The next step in the competence regulatory cascade involves two other early gene products: ComX (a.k.a. SigX or σ^X) and ComW. When these two proteins are expressed, ComW exchanges the housekeeping sigma factor, σ^A, in RNA polymerase for ComX (94). ComX directs RNA polymerase to bind to an alternative −10 element known as the cin-box in the promoters of the so-called late (competence) genes and upregulate their expression (95, 96). A subset of these late genes encodes the machinery required for transformation (97). Their expression allows transformation to occur and marks entry into the bona fide competence state.

The organization of the *com* system leads to certain characteristic features of the timing of pneumococcal competence activation. First, because *com* enters a positive feedback loop upon activation, the ensuing massive increase in CSP signaling tends to trigger entry into competence simultaneously for almost all cells within a population (98). Second, the dual-stage nature of the regulatory cascade, with CSP signaling first inducing early gene expression through ComE, and subsequently late gene expression through ComX, results in a five- to ten-minute lag between when CSP first activates the system and cells becoming transformable (28). While ComE responds almost immediately to CSP to induce early gene expression, it takes five minutes for ComX and

ComW levels to increase sufficiently to trigger late gene expression to render the cells transformable (93, 97, 99).

There exists some genetic diversity among different strains within the *com* system. The major source of diversity is found in ComC and ComD. Any one strain can encode one of many alleles, or pherotypes, of ComC, and one of many alleles of ComD, which is usually correlated with its ComC pherotype (100). Most strains encode one of two major ComC pherotypes, named ComC-1 and ComC-2 (CSP-1 and CSP-2). There is limited cross-stimulation between non-cognate CSP-ComD pairs (101), and strains expressing one pherotype will typically not cross-stimulate strains expressing a different pherotype. As a result, pherotype diversity allows different populations to activate competence independently of one another even if they are mixed in co-culture. Besides pherotype diversity, *com* is nearly universally conserved in pneumococcus; various loss-of-function mutations are found in only a handful of strains (102).

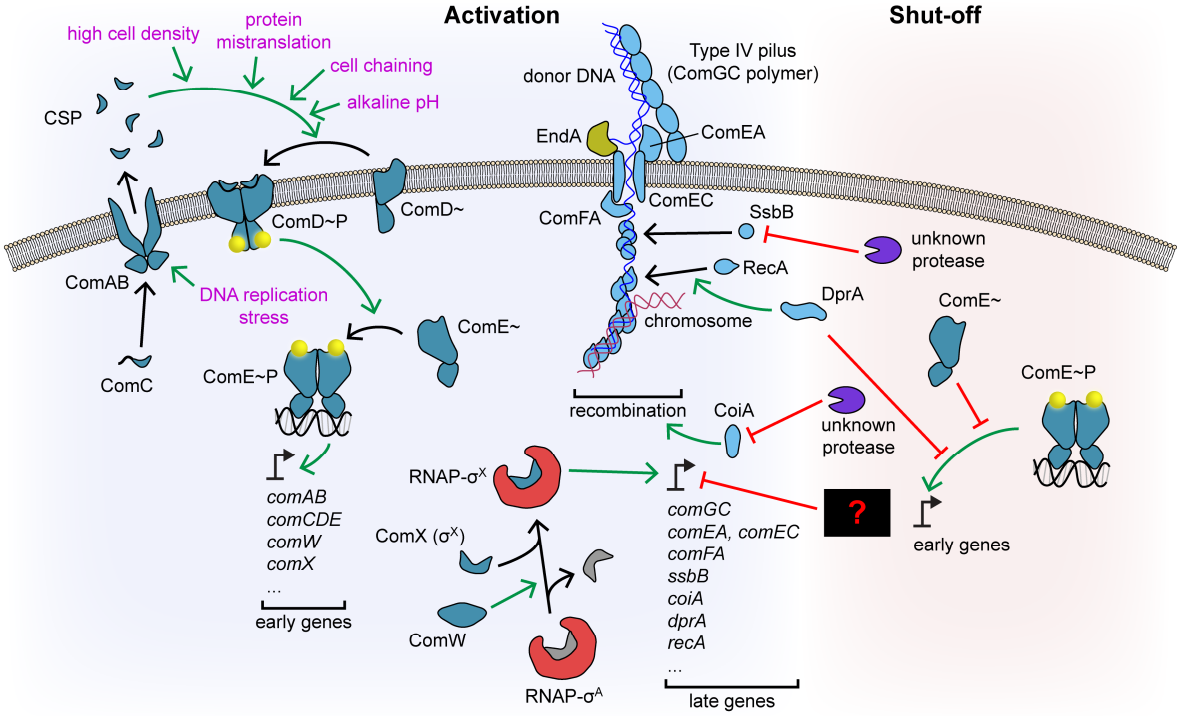

Figure 1.4. Competence regulation in pneumococcus. Early competence gene products are colored dark blue. Late competence gene products are colored light blue. X~ denotes unphosphorylated form of X. X~P denotes phosphorylated form of X. RNAP: RNA polymerase.

The *com* system was originally described as a classical quorum-sensing system, with CSP acting as a freely diffusible signal allowing cell-cell communication over distance. However, recent studies have suggested that CSP diffusion is more limited than previously thought. Prudhomme et al. showed that CSP does not freely diffuse throughout a liquid culture but instead is retained on the surface of cells (103). Consequently, the timing of competence activation is controlled primarily by time of growth since inoculation instead of cell density. This led the authors to conclude that *com* functions as a timing mechanism where activation occurs stochastically in individual cells and is spread across a population through cell contact. A later study contradicted these results, finding that CSP could in fact diffuse into the supernatant of broth cultures and that the mode of *com* activation is consistent with extracellular CSP concentration (which is correlated imperfectly with cell density) reaching a critical threshold rather than a timing mechanism (104). The authors of the second study found that CSP retention on cells differed between encapsulated and non-encapsulated strains and offered this as a partial explanation for the conflicting results. The natural setting for competence development is within biofilms on the nasopharyngeal mucosa (105) where both cellular and pheromone diffusion is likely to be limited. Therefore, the details of CSP diffusion and how this affects propagation of signal across a structured population is a crucial open question in the field.

While the ComABCDE complex forms the core regulatory loop that ultimately controls when pneumococcus becomes competent, many inputs modulate the loop to promote or inhibit its activation (Fig. 1.4, magenta text). Examples of such inputs include cell density, pH, temperature (106, 107), antibiotic stress (108, 109), and protein mistranslation (110). All these inputs must directly or indirectly affect some aspect of the core regulatory loop. Cell density is directly correlated with extracellular CSP concentrations; more secreting cells leads to greater CSP accumulation. In this way increasing cell density promotes competence activation in a classical quorum-sensing fashion. The pH of the growth medium affects regulation by modulating the binding and response to CSP of ComD (111). DNA synthesis-inhibiting antibiotics such as fluoroquinolones increase the dosage of genes near the origin of replication, including *comAB*, which in turn increases ComAB expression and CSP secretion (108). Antibiotic

stress can also induce increased cell chaining, which increases the local concentration of CSP around chains (109). Protein mistranslation shifts the activity of the multi-functional extracellular protease HtrA toward protein quality control and away from CSP degradation, increasing the rate of extracellular CSP accumulation (110, 112). In addition to the above, the CiaRH two-component system inhibits competence in two ways: 1) inducing the expression of short, non-coding RNAs which interfere with ComC translation and 2) inducing the expression of the aforementioned CSP-degrading HtrA (113, 114). While CiaRH is known to be activated following competence activation (115), the specific activating signal for the system remains undefined, leaving the purpose of CiaRH-regulated competence inhibition a mystery for now. Even if some of them are not yet fully understood, the many inputs integrated by *com* to adjust competence regulation underscore the great lengths to which pneumococcus has gone to finely tune exactly when and where competence occurs.

1.3.2 Competence Shut-Off

An intriguing feature of competence in *S. pneumoniae* is that once activated, competence is relatively short-lived. Exogenous CSP-induced competence lasts 30 to 40 minutes (28), and episodes of natural, spontaneous competence activation in broth culture last slightly longer at around one hour (107). This rapid shut-off of competence is also reflected in the transcriptional kinetics of early and late genes. Early gene transcript levels peak five minutes after CSP treatment, decrease substantially at 10 minutes, and return almost to baseline by 20 minutes; late gene transcript levels peak 10 minutes after CSP treatment and return to baseline by 20-30 minutes (88, 93, 97, 99, 116). Moreover, during the shut-off phase both transformability and early/late gene transcription become refractory to CSP signaling (88, 117). Therefore, shut-off cannot be solely explained by degradation of extracellular CSP and subsequent loss of signaling.

The current model of competence shut-off incorporates multiple related but distinct mechanisms that work in concert (Fig. 1.4, right). First, at least some of the late gene products required for transformation are unstable and disappear from cells at roughly the same time that competence shuts off. Two examples are SsbB (118), the competence-specific single-stranded DNA binding protein required to stabilize donor

DNA (119), and CoiA (120), a protein required for efficient recombination and integration of donor DNA into the chromosome (119). It is unclear if all components of the transformation machinery are similarly unstable, but since each individual component is required for transformation, the timely degradation of even some of them is sufficient to explain why cells rapidly exit from competence. However, for these components to be cleared from the cell effectively, their expression needs to be stopped. Therefore, the cessation of early and late gene expression are key parts of competence shut-off and need to be explained.

A second mechanism helps to explain the cessation of early gene expression. Unlike phospho-ComE (ComE~P) which promotes early gene transcription, unphosphorylated ComE (ComE~) binds to the promoters of early genes and inhibits their transcription (89, 121). During the early stages of *com* activation, ComD-mediated phosphorylation supplies enough ComE~P to outcompete ComE~, resulting in increased early gene expression. However, because ComE~P also induces expression of ComE, over time newly synthesized ComE~ builds up as the phosphorylation capacity of ComD struggles to keep pace. This leads to ComE~ displacing ComE~P from early gene promoters and eventual shut-off of early gene transcription. This mechanism also explains why *com* is refractory to CSP signaling during shut-off; ComE~P is responsible for directly promoting early gene transcription and indirectly promoting late gene transcription through ComX/ComW in response to CSP but cannot perform this task due to being outcompeted by ComE~.

A third mechanism is mediated by the late-gene product DprA. The primary function of DprA is to load the recombinase RecA onto donor DNA during transformation (73). Recently DprA was also shown to bind ComE~P and prevent its binding to DNA, either through direct disruption of DNA binding, disruption of dimer formation, or sequestration (118). This potentiates the effect of the second mechanism described above by removing ComE~P from circulation and further biasing the ratio of ComE~ to ComE~P in favor of ComE~.

The above two mechanisms together account for the shut-off of early gene transcription and the CSP-refractory period, but do not by themselves explain the rapid shut-off of late gene transcription. Decoupling ComX and ComW from ComE regulation

results in a slightly delayed but still rapid shut-off of competence (122), indicating that something separate from ComE also contributes to competence shut-off. The observation that ComX rapidly disappears from cells following competence activation (96) suggested that perhaps degradation of ComX leads to shut-off of late gene expression. However, prolonging the half-lives of both ComX and ComW by inactivating the Clp protease responsible for their degradation did not prolong either late gene expression or competence (123). Therefore, a non-proteolytic mechanism mediates the inhibition of ComX activity that leads to late gene expression shut-off. So far, the nature of this inhibition and the factors responsible for it have eluded detection.

1.3.3 Interbacterial DNA Predation

Once *S. pneumoniae* becomes competent, it is ready to take up extracellular donor DNA for transformation. However, obtaining access to a source of donor DNA presents several issues. First, the current model of *com* regulation does not provide any means by which cells can sense the presence of extracellular DNA and preferentially activate competence under these conditions. Second, because the duration of competence once activated is so short, it is unlikely that cells will encounter extracellular DNA during this brief period by random chance alone.

To increase its chances of obtaining donor DNA during competence, pneumococcus has evolved strategies to liberate DNA from neighboring cells through lysis to use for transformation in a process called fratricide. Fratricide has been shown to kill both pneumococcus and closely related Mitis group streptococci (124). During competence, pneumococcus upregulates the expression of three late gene products: CbpD, LytA, and CibAB (74, 116). CbpD is a murein hydrolase that degrades the cell wall of target cells (125). LytA is an amidase that also degrades the cell wall and promotes self-lysis (autolysis) (126). CibAB is a two-peptide bacteriocin made up of the processed and secreted forms of two GG peptides, CibA and CibB (127). In addition to these, the constitutively expressed autolytic lysozyme LytC (128) also participates in fratricide. Collectively, these factors are called fratricide effectors or fratricins.

During fratricide, CbpD and CibAB are released by competent cells to target non-competent neighbors and trigger the action of the autolysins LytA and LytC, which are required for maximal killing efficiency (127, 129). Competent cells are protected from

their own CbpD and CibAB by the immunity proteins ComM and CibC, respectively, which are specifically expressed during competence (127, 130). While CbpD and CibAB are exclusively produced by the killer (competent) cells, LytA and LytC can be supplied by either the killer or the target cells; in fact target-derived LytA and LytC contribute more to fratricide than their killer-derived counterparts (127, 129).

None of the fratricide effectors are required for transformation, and fratricide-deficient strains transform as efficiently as their wild-type counterparts when free, purified donor DNA is supplied (124). On the other hand, fratricide increases transformation efficiency by 10- to 1000-fold when competent cells must obtain donor DNA from live, intact cells in liquid co-culture (124). Fratricide also increases transformation efficiency by a comparable factor (50-fold) in pneumococcal biofilms (131). These data indicate that pneumococcus has evolved specifically to exploit neighboring cells as a source of donor DNA to use during competence and fratricide is the key means by which it taps into that source.

While fratricide is certainly an important tool for DNA acquisition, the large increases in transformation efficiency owed to fratricide reported through in vitro experiments risk overselling its effectiveness. In the previously mentioned studies (124, 131), the recipient/donor pair was usually chosen such that the donor strain could not protect itself from fratricide. In reality, the bacteria targeted by pneumococcal fratricide, namely pneumococcus itself and other Mitis group streptococci, normally encode the appropriate factors that confer immunity to fratricide. These factors are exclusively expressed during competence, so potential donor cells only need to activate competence to become immune. Moreover, any potential donor cell close enough to the would-be recipient to be killed by fratricide is also close enough to sense CSP signaling from the latter to activate competence. For donor cells to remain susceptible to fratricide then, they must either belong to the small subset of cells (<2%) in a population that does not respond normally to CSP signaling (127, 132) or belong to a different CSP pherogroup from the recipient cells. Indeed, genomic analyses suggest that genetic exchange between pneumococcal CSP pherogroups may be more common than within pherogroups (133). However, since nearly all pneumococci belong to one of just two CSP pherogroups (100), this does not provide many chances for cross-group fratricide

and DNA exchange among pneumococci. Because CSP pherotype diversity is much greater in non-pneumococcal Mitis group streptococci (134), cross-species fratricide does not face the same barrier.

Perhaps to overcome the limitations of fratricide discussed above, pneumococcus has evolved to augment fratricide with bacteriocin-mediated killing. In pneumococcus, activation of the *com* system also induces the production of the Blp bacteriocins (39, 40). This occurs through two mechanisms: 1) the ComAB transporter secretes the peptide pheromone, BlpC, that activates the quorum-sensing system controlling Blp bacteriocin production, and 2) ComE directly upregulates the expression of BlpC and its native transporter, BlpAB. Preventing communication between *com* and the Blp system impairs genetic exchange from a Blp bacteriocin-sensitive strain to a Blp bacteriocin-producing strain in a biofilm (40). Crucially, in this experiment both strains could become competent and responded to the same CSP pherotype. Therefore, DNA acquisition through Blp bacteriocin-mediated killing can function in the absence of efficient fratricide. Because the Blp bacteriocins and their immunity factors are regulated by the same quorum-sensing system, BlpC signaling can make potential donor cells immune to Blp bacteriocins similarly to how CSP signaling can make them immune to fratricide. However, BlpC pherotype diversity is greater than that of CSP (135, 136). Additionally, unlike the fratricide effectors and immunity factors which are conserved among pneumococci, Blp bacteriocin and immunity factor content can vary from strain to strain (137, 138). Therefore, not only can two strains belong to different BlpC pherogroups and be unable to cross-activate each other's Blp systems, they can also produce and be immune to completely different bacteriocins. These properties of Blp increase the chances of a competent, Blp bacteriocin-producing strain encountering another strain susceptible to killing by the former's bacteriocins and therefore a viable source of accessible donor DNA.

1.3.4 Functional Significance of Competence

Genetic competence in general has been proposed to fill three primary roles: nutritional intake, DNA repair and genome maintenance, and horizontal gene transfer (HGT) (26). Of these three, the latter two have compelling evidence supporting them as functions of competence in *S. pneumoniae*.

In the case of DNA repair, it has been shown that DNA-damaging stress induces competence in pneumococcus (139). Competence activation is also positively correlated with increasing chromosomal mutation burden (140). Transformation by homologous recombination is expected to remove damaged or mutated DNA by providing a means for mutants to re-acquire the original sequence from non-mutants among the rest of the population. In support of this model, competence in pneumococcus has been shown to reduce the fixation rate of new mutations in laboratory conditions (141). Therefore, pneumococcal competence likely helps to preserve genome integrity.

While the above processes tend to reduce or stabilize genetic diversity within a population, transformation during competence has also been well documented as a source of *increasing* genetic diversity. This is accomplished through HGT, in which pieces of DNA are passed among different bacteria. HGT has been documented in pneumococcus in diverse settings: within a single host during chronic infection (142), among strains in a geographically isolated population (143), and in a global, pandemic lineage (79). The primary site where transformation occurs in pneumococcus is likely its natural habitat, the human nasopharynx (105). Up to 50% of children who are colonized with pneumococcus are colonized with multiple strains (47). Therefore, colonizing pneumococci often can exploit within-species genetic diversity for HGT during competence. Moreover, a variety of non-pneumococcal streptococci species also inhabit the nasopharynx, including closely related members of the Mitis group (45, 46). As a result, pneumococcus can also use competence to facilitate cross-species genetic exchange. Numerous in vitro experiments have demonstrated that pneumococcus can indeed acquire DNA from other streptococci via transformation (124, 144-146). There is also ample genomic evidence of naturally occurring genetic transfer from other streptococci, most notably *S. mitis*, to pneumococcus via recombination (134, 147-149). This indicates that competence allows pneumococcus to use not only other strains but also other streptococcal species as reservoirs of genetic diversity.

HGT through transformation has shaped pneumococcus's adaptive response to antibiotics and vaccination. Genetic determinants of antibiotic resistance are hotspots of recombination (143). In particular, Transformation and recombination at the *pbp* locus

has been shown to facilitate the spread of beta lactam resistance among pneumococci (80, 143, 150-152). Cross-species transformation has allowed pneumococcus to acquire beta lactam resistance from other Mitis group streptococci as well (148, 149). Serotype switching occurs via recombination at the capsule biosynthesis locus *cps*. Since this locus is rather large (10 to 30 kb) (153), recombination events large enough to result in a serotype switch are predicted to be rare. However, the observed frequency of serotype switching is higher than expected, suggesting that selection is acting on these events (79). Since the introduction of PCV, multiple cases of vaccine-serotype strains switching to non-vaccine types via recombination have been observed, a process called vaccine escape. In the United States, serotype 4 (included in PCV7) strains switched to serotype 19A (not included in PCV7) in five separate instances, one of which resulted in a highly successful vaccine escape strain that quickly spread across the country (81, 154). In another example, the multidrug-resistant global PMEN14 lineage was originally found with serotype 19F (included in PCV7) but in four separate instances switched to serotype 19A (80). The introduction of PCV13, which includes protection against serotype 19A, has begun to reverse the expansion of 19A strains (155). However, history is repeating itself as serotype switching from the vaccine types 9V, 14, 19A, and 19F to the non-PCV13 type 35B has created new multidrug-resistant vaccine escape strains (67). These findings underscore the urgent need for further investigation of competence and transformation. Doing so will improve our understanding of how these processes shape the way the pneumococcus responds to efforts to treat and prevent pneumococcal disease.

1.4 Bacteriocin Production in *Streptococcus pneumoniae*

S. pneumoniae encodes a diverse set of bacteriocins. Currently, 21 different pneumococcus-encoded bacteriocin clusters have been identified based on sequence analysis (156). Pneumococcus's bacteriocin repertoire spans at least six bacteriocin families and includes 11 lantibiotics (157) (streptolancidins), five lactococcin 972 (158)-like bacteriocins (streptococcins), one class IIc (159) circular bacteriocin, one lassopeptide (160), one sactipeptide (161), the competence-induced bacteriocin CibAB, and the Blp bacteriocins.

The epidemiology of pneumococcal bacteriocins varies widely among different clusters (156). Some, such as Blp and CibAB, are present in all or nearly all strains. Others, such as streptolancidin A, are only found in a handful or even a single strain. Moreover, some clusters, most notably Blp, exhibit great diversity among different strains. Very few clusters have been characterized experimentally; for the vast majority little to nothing is known about their regulation, activity, spectra of inhibition, or modes of action. Knowledge about the phenotypic consequences of genotypic strain-to-strain differences within clusters is equally lacking.

Although much needs to be done to characterize the pneumococcal bacteriocins, a select few have been studied in detail. These include CibAB, several streptolancidins, and the Blp bacteriocins. The following sections will summarize what has been discovered about these bacteriocins.

1.4.1 The Competence-Induced Bacteriocin CibAB

The competence-induced bacteriocin CibAB is a class IIb two-peptide bacteriocin produced exclusively during competence (116, 127). The primary mode of action of class IIb bacteriocins is proposed to be formation of ion-selective membrane pores (162). The component peptides of CibAB, CibA and CibB, are produced initially as double-glycine precursor peptides and then presumably processed and secreted by the competence-related PCAT ComAB. Like with all two-peptide bacteriocins (162), both CibA and CibB are required for optimal killing activity (127). Immunity to CibAB is conferred by the co-expressed protein CibC. The initial characterization of CibAB found that it was active against pneumococcus, but it only retained killing activity in the presence of the autolysins LytA and LytC (127). The original function ascribed to CibAB was to kill neighboring cells during competence to gain access to their DNA for transformation in a process called fratricide. For a detailed discussion of fratricide, please refer to the previous sub-chapter.

Another role for CibAB involving colonization of the nasopharynx was recently discovered (163). Interestingly, while this new role requires competence activation, it is unrelated to transformation. Once pneumococcus (the resident) establishes itself in the nasopharynx, it inhibits colonization of the same host by other strains (the invaders). CibAB is one of two factors that mediate this competitive inhibition (the other is fellow

fratricide effector CbpD). It is not known if CibAB-mediated competitive inhibition of colonization requires LytA and LytC like CibAB-mediated fratricide does. LytA- and LytC-null residents were shown to have no defect in competitive inhibition. However, it was previously established for fratricide that LytA and LytC could be provided by the target cell (127). Therefore, for competitive colonization it is possible that LytA and LytC derived from invader cells are sufficient to potentiate CibAB-mediated inhibition. This is one of many questions still to be answered about this newly discovered role for bacteriocins in pneumococcus.

1.4.2 The Streptolancidins

The streptolancidins belong to the lantibiotic (lanthionine-containing antibiotic) family of bacteriocins. Lantibiotics are extensively post-translationally modified and contain the characteristic cyclized thioether amino acids lanthionine and methyllanthionine in addition to other unusual modifications (157). Lantibiotics kill their targets by inhibiting cell wall synthesis, either alone or in combination with membrane pore formation (164). To prevent self-killing by producer strains, lantibiotic clusters contain factors that confer immunity to their respective lantibiotics. Two types of immunity factors have been discovered: LanFEG and LanI. LanFEG proteins are ABC transporters that likely contribute to immunity by secreting lantibiotics out of the cell, while LanI proteins confer immunity through an unknown mechanism (157).

Lantibiotics are initially synthesized as inactive precursor peptides with N-terminal leader peptides (signal sequences). Modification of the precursor peptides occurs with the leader peptide still attached. Formation of the (methyl)lanthionine rings is catalyzed by LanB and LanC for class I lantibiotics, LanM for class II, LanKC for class III, and LanL for class IV (165). Other modification enzymes, such as the decarboxylase LanD, are present in certain clusters (166). The protease LanP removes the leader peptides of modified precursors and the mature core peptides are then secreted by the transporter LanT. As an exception to the above rule, class II lantibiotics are double-glycine peptides and therefore have their leader peptides removed by a PCAT-type LanT concomitantly with secretion instead of relying on LanP for leader peptide removal (157).

The first lantibiotic cluster in pneumococcus to be experimentally characterized was streptolancidin G. The streptolancidin G cluster (*slg*) is a putative class II lantibiotic cluster containing two precursor peptides, modification enzymes SlgM and SlgD, and the PCAT SlgT (167). Even though SlgT is expected to catalyze leader peptide removal by itself, *slg* also contains two putative LanP-type proteases, SlgP1 and SlgP2. The exact roles SlgT and SlgP1/2 play in leader peptide removal remain unclear. This cluster was found to be regulated by a peptide pheromone signaling system named TprA/PhrA (168). In this system, the Phr-family pheromone PhrA is secreted and then re-imported into the cell by the Ami/Opp oligopeptide importer. After re-importation, PhrA antagonizes the activity of the PlcR-family transcriptional repressor TprA. TprA normally represses transcription of *slg*. PhrA signaling therefore acts to derepress *slg* and induce expression of the streptolancidin G biosynthesis machinery. PhrA signaling is inhibited in glucose-containing media due to catabolite repression. As a result, *slg* induction requires growth with an alternative carbon source such as galactose. A homologous signaling system found in certain lineages, TprA2/PhrA2, which regulates another streptolancidin cluster, has been shown to cross-activate the TprA/PhrA system (169). While its regulation is well-characterized, streptolancidin G has not been shown to exhibit antimicrobial activity in pneumococcus. The *slg* cluster lacks any LanI or LanFEG immunity proteins, suggesting perhaps the peptides are inactive. On the other hand, in vitro modification of the streptolancidin G core peptides by the heterologous class I nisin modification enzymes yielded modified peptides with antimicrobial activity (167). This shows that the streptolancidin G precursors can be modified to an active form. However, it remains to be seen whether the biosynthetic enzymes encoded by *slg* can effect all the necessary modifications to convert streptolancidin G to its active form.

To date, the only pneumococcal lantibiotic cluster confirmed to produce antimicrobial activity is streptolancidin A, which has been found intact in only one strain (156). The Streptolancidin A cluster (*sla*) encodes four similar class II precursor peptides, SlaA1-4, modification enzyme SlaM, transporter SlaT, a two-component regulatory system SlaKR, and an immunity ABC transporter SlaFE (170). The lantibiotic peptides serve dual roles as both effectors and signaling pheromones. The cluster is regulated by an autocatalytic positive feedback loop in which the mature peptides signal

through SlaKR to upregulate their own production. A unique feature of *sla* is the presence of four distinct but structurally similar lantibiotic peptides. Of these, one (SlaA4) is dispensable for both signaling and antimicrobial activity. The remaining three peptides must work together to upregulate *sla* and to kill target cells but their exact roles in these processes remain unknown. Streptolancidin A shows broad-spectrum antimicrobial activity in vitro against *S. pneumoniae*, various other streptococci, *L. lactis*, and *Listeria monocytogenes*. This in vitro activity also translates to an in vivo setting; in a mouse model, production of streptolancidin A confers a competitive advantage to producer strains over sensitive strains during nasopharyngeal colonization (170).

1.4.3 The Blp Bacteriocins (Pneumocins)

The Blp bacteriocins (pneumocins) are a diverse group of unmodified, single- and two-peptide bacteriocins (138, 171-173) and are the most well-studied of the pneumococcal bacteriocins. Pneumocins are encoded as double-glycine precursor peptides in the *blp* locus, which is present in all sequenced pneumococcal strains (156). The *blp* locus is organized into two halves. The first half contains genes involved in regulation (*blpRH*, *blpC*) and transport (*blpAB*). The second half, called the bacteriocin-immunity region (BIR), contains the genes encoding the pneumocin precursor peptides and their immunity proteins.

A quorum-sensing system similar to *com*, which regulates competence, controls activation of *blp* (24). The PCAT BlpAB secretes BlpC, a double-glycine peptide pheromone (14). When extracellular BlpC reaches a threshold concentration, it signals through the BlpHR two-component system to promote transcription of target genes (24). When stimulated, the response regulator BlpR acts as a transcriptional activator to directly enhance expression of *blpABC*, completing a positive feedback loop. BlpR also activates expression of the pneumocin precursors and immunity proteins from the BIR. Upon expression, the pneumocin precursors are processed to their active forms and secreted by BlpAB (14). As a result, classical *blp* activation simultaneously produces active pneumocins and renders the producers immune to their own pneumocins.

While some version of *blp* is found in all pneumococci, the content of the locus can vary greatly among different strains. The region with greatest diversity is the BIR. Twelve unique putative *blp* bacteriocin genes have been identified. Any single strain’s

BIR encodes between zero and seven of these bacteriocin peptides, between zero and nine putative immunity proteins, and various accessory proteins of unknown function (138). Most BIR genes have more than one allelic variant, which also contributes to diversity. Outside of the BIR, variation is found in *blpH* and *blpC* in the form of pherotype diversity. Twenty different allelic variants of the mature BlpC pheromone have been found, with nine "common" pherotypes appearing in at least 0.5% of strains (136, 138). Each BlpC pherotype is usually paired with a matching BlpH receptor variant, and there is limited cross-talk among the different pherotypes (135, 136).

Strain-to-strain differences in pneumocin and immunity protein content in conjunction with pherotype diversity are thought to promote competition among distantly related strains while promoting cooperation or neutrality between clonal and closely related strains. Two strains with mismatched pneumocins/immunity proteins can kill each other trivially because neither one can become immune to the other's pneumocins. Two strains with mismatched BlpC pherotypes can kill each other even if they have matched pneumocins/immunity proteins because they can activate *blp* at different times; since pneumocin immunity requires *blp* activation this creates the possibility of one strain producing pneumocins during a time when the other is not immune. In contrast, strains with matching BlpC pherotypes and pneumocins/immunity proteins will cross-activate each other's *blp* systems and thus are always protected from being killed by each other's pneumocins.

Another common source of diversity in *blp* is in the genes encoding the transporter BlpAB. The *blpA* gene encodes the PCAT proper, and the *blpB* gene encodes an accessory protein of unknown function which is required for transport. Roughly 75% of strains have mutations in *blpA* and/or *blpB* that result in the production of a non-functional transporter (14, 138). Because BlpAB secretes both BlpC and the pneumocins, it was initially thought that strains with non-functional BlpAB ($BlpAB^{-}$) could neither produce active pneumocins nor activate *blp* by themselves. However, most $BlpAB^{-}$ strains can still sense exogenously provided BlpC and synthesize their pneumocin immunity proteins in response. Therefore, these strains were dubbed "cheaters" which in the presence of $BlpAB^{+}$ "producer" strains could protect themselves from pneumocin-mediated killing and reap the benefits of the elimination of pneumocin-

sensitive strains without having to expend the resources to produce pneumocins themselves (14).

The producer-cheater dichotomy has recently been challenged, however. The *com* transporter ComAB – another PCAT with a high degree of similarity to BlpAB – was found to secrete BlpC, and the *com* response regulator ComE was found to directly upregulate *blpABC* (39, 40). This crosstalk allows competence activation to stimulate *blp* activation independently of BlpAB. In BlpAB$^+$ strains cross-talk allows pneumocins to contribute to DNA acquisition during competence (40). In BlpAB$^-$ strains crosstalk allows self-activation of *blp*, albeit only during competence. Additionally, though experimentally unverified, it is expected that ComAB can secrete the pneumocins in addition to BlpC due to similarities in their signal sequences. This would theoretically allow BlpAB$^-$ strains to also produce active pneumocins during *blp* activation. While Lux et al. (137) observed pneumocin-mediated inhibition by BlpAB$^-$ strains, others have found either no or substantially reduced inhibition by BlpAB$^-$ strains compared to BlpAB$^+$ strains (14, 173, 174). These conflicting results are potentially explained by differences in competence activation between assays, but it remains to be seen exactly how the discrepancies are resolved.

Of the twelve identified putative pneumocin genes, five have been confirmed to code for active bacteriocins. The *blpM* and *blpN* genes encode the two-peptide bacteriocin BlpMN, *blpI* and *blpJ* encode another two-peptide bacteriocin BlpIJ, and *blpK* encodes the single-peptide bacteriocin BlpK (171, 173). Pneumocins show in vitro antimicrobial activity against pneumococcus and other Mitis group streptococci such as *S. mitis*, *S. oralis*, and *S. salivarius*, as well as the more distantly related *Streptococcus pyogenes*, *Lactococcus lactis*, and *Micrococcus luteus* (137, 173). In vitro biofilm and mouse nasopharyngeal colonization models reveal a competitive advantage for pneumocin-producing strains over pneumocin-sensitive strains (14, 40, 171, 173). However, an effort to extend this phenotype to colonization of humans failed to show a role for *blp* in restricting co-colonization by strains predicted to compete via pneumocin-mediated inhibition (174). The authors of this study present several possible explanations for this discrepancy: 1) their methods were insensitive to certain types of competition (for example, where one strain reduces the colonization density of another

without completely eliminating it, or where colonization is reduced in duration instead of outright prevented), 2) pneumocin-mediated inhibition is restricted to certain conditions, or 3) the primary role of pneumocins is something other than competitive inhibition (for example, DNA acquisition during competence). It is up to future studies to determine which of these, if any, is true and to discover the details of how the pneumocins shape the colonization patterns of and interactions between pneumococci in humans.

1.5 Outline of the Thesis

The central goal of this thesis is to investigate the regulation and substrate pools of three PCATs found in *S. pneumoniae* to determine how each one contributes to genetic competence and bacteriocin production. In Chapter II, I examine the contributions of the competence transporter ComAB and the pneumocin transporter BlpAB to competence regulation and pneumocin production. I find that ComAB and BlpAB share the same substrate pool, which mediates bi-directional crosstalk between the *com* and *blp* regulatory systems. I also show that differences in peptide pheromone and pneumocin secretion between $BlpAB^{+}$ and $BlpAB^{-}$ strains separate them into aggressor and opportunist pneumocin users, respectively. In Chapter III, I characterize *rtg*, a previously undescribed putative bacteriocin locus in *S. pneumoniae* encoding a PCAT, RtgAB, and multiple double-glycine peptides. I describe the regulation of *rtg* and the contribution of *rtg* to inter-bacterial competition. I also determine that RtgAB secretes a different set of peptides from ComAB and BlpAB, and I explore the features within the peptide substrates underlying this selectivity. In Chapter IV, I summarize and discuss the major findings of this thesis as well as important unanswered questions related to this work.

Chapter II: ABC Transporter Content Diversity in *Streptococcus pneumoniae* Impacts Competence Regulation and Bacteriocin Production.

2.1 Abstract

The opportunistic pathogen *Streptococcus pneumoniae* (pneumococcus) uses natural genetic competence to increase its adaptability through horizontal gene transfer. One method of acquiring DNA is through predation of neighboring strains with antimicrobial peptides called bacteriocins. Competence and production of the major family of pneumococcal bacteriocins, pneumocins, are regulated by the quorum sensing systems *com* and *blp*, respectively. In the classical paradigm, the ABC transporters ComAB and BlpAB each secrete their own system's signaling pheromone and in the case of BlpAB also the pneumocins. While ComAB is found in all pneumococci, only 25% of strains encode an intact version of BlpAB (BlpAB$^+$) while the rest do not (BlpAB$^-$). Contrary to the classical paradigm, it was previously shown that BlpAB$^-$ strains can activate *blp* through ComAB-mediated secretion of the *blp* pheromone during brief periods of competence. To better understand the full extent of *com-blp* crosstalk, we examined the contribution of each transporter to competence development and pneumocin secretion. We found that BlpAB$^+$ strains have a greater capacity for competence activation through BlpAB-mediated secretion of the *com* pheromone. Similarly, we show that ComAB and BlpAB are promiscuous and can both secrete pneumocins. Consequently, differences in pneumocin secretion between BlpAB$^+$ and BlpAB$^-$ strains derive from the regulation and kinetics of transporter expression rather than substrate specificity. We speculate that BlpAB$^-$ strains (opportunists) use pneumocins mainly in a narrowly tailored role for DNA acquisition and defense during competence while BlpAB$^+$ strains (aggressors) expand their use for the general inhibition of rival strains.

2.2 Introduction

Streptococcus pneumoniae (pneumococcus) is an opportunistic pathogen that can cause serious illnesses such as pneumonia, meningitis, and bacteremia, with the greatest disease burden in the very young and the elderly. The natural niche of pneumococcus is the human nasopharynx, and colonization of this niche is a prerequisite for invasive pneumococcal disease. Pneumococcus colonizes up to 60% of young children (175, 176). As many as half of those who are colonized carry multiple pneumococcal strains (47).

Pneumococcus, a naturally competent bacterium (68), can exploit the large pool of genetic material available to it (175-177) in the nasopharynx. Natural competence allows pneumococcus to take up new genetic material through horizontal gene transfer and recombination. Multiple studies have documented that recombination occurs with great frequency in pneumococcal lineages that are globally distributed (79), geographically isolated (143), and even confined to a single patient (142). Additionally, to compete with other bacteria found in the nasopharynx, pneumococcus produces small anti-microbial peptides called bacteriocins. The major family of bacteriocins encoded by pneumococcus are the pneumocins. The pneumocin locus, *blp*, is found in all sequenced strains of pneumococcus (172). Pneumocin-producing organisms inhibit sensitive strains and have a fitness advantage in both in vitro biofilms and competitive colonization of the mouse nasopharynx (40, 171).

Both competence and pneumocin production in pneumococcus are under strict regulation by two separate but similar systems (Fig. 2.1). The *com* system regulates competence. In this system, a peptide pre-pheromone, ComC, is processed and secreted by a transporter complex ComAB (27, 28). After processing and secretion, the mature pheromone, now called competence-stimulating peptide (CSP), accumulates extracellularly. Once a threshold concentration is reached, CSP signals through the ComDE two-component system to upregulate the set of so-called “early (competence) genes” (87). The early genes include *comAB* and *comCDE*, creating a positive feedback loop. Upregulation of early gene expression starts a regulatory cascade mediated by the alternative sigma factor ComX that ultimately leads to competence development (178). The *com* system integrates many environmental and physiological signals, such as cell

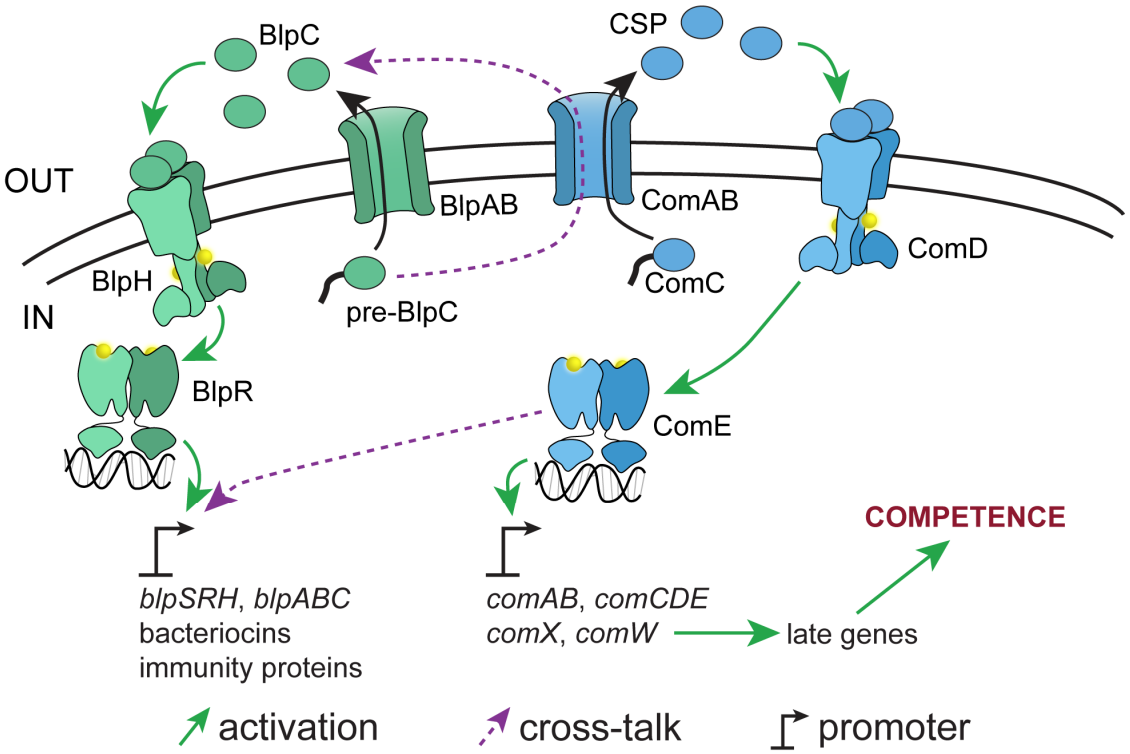

Figure 2.1. *com* and *blp* regulation.

density (106), pH (107), antibiotic stress (108) and protein mistranslation (110). As a result, the propensity for competence activation can differ greatly from one set of conditions to another.

Meanwhile, the *blp* locus regulates pneumocin production in a manner similar to *com* and competence (179). In the prototypical case, a small peptide pheromone, BlpC, is processed and secreted by the BlpAB transporter complex. Mature BlpC then signals through the BlpHR two-component system to upregulate the entire *blp* locus. Unlike with *com*, the entire *blp* regulon is directly controlled by BlpR. The upregulation of the regulatory system forms another positive feedback loop, while the upregulation of the so-called bacteriocin immunity region (BIR) within the *blp* locus results in the production of a diverse array of pneumocins and their immunity proteins. Like with competence, the propensity for *blp* activation is environment and context-dependent.

Due to the auto-inducing nature of *com* and *blp*, activation tends to proceed synchronously among all cells within a population once the pheromone concentration threshold is reached (28). However, different strains can encode different pherotypes of

CSP and BlpC along with matched cognate ComD and BlpH receptors (101, 135). In general, each pherotype only efficiently activates its cognate receptor type. Therefore, CSP/BlpC signaling and synchronous *com*/*blp* activation is restricted by pherotype. Pherotype diversity among pneumococci may have evolved as a method for cells to privilege clonal or closely-related cells which are more likely to have a matched pherotype. Such cells would then share in the benefits of competence activation or be protected from pneumocin-mediated killing while strains with mismatched pherotypes would not.

While *com* and *blp* were originally thought to operate independently, two recent studies have shown that *com* positively influences the regulation of *blp* (39, 40). This occurs through two mechanisms (Fig. 2.1): 1) ComE directly upregulates transcription of *blp* genes, and 2) ComAB processes and secretes BlpC in addition to ComC/CSP. As a result, *com* activation also induces *blp* activation. This crosstalk from *com* to *blp* allows for the coordinated regulation of the two systems. In a biofilm model, pneumocin production resulted in an increase in transformation efficiency (40). The competence regulon itself encodes so-called fratricide effectors which target and kill non-competent pneumococci (180). The role of fratricide seems to be to increase access to extracellular DNA during competence (124, 131). Therefore, pneumococcus may have evolved *com* to *blp* crosstalk to increase access to genetic material during competence by using pneumocins to augment fratricide.

While the *com* regulatory system is found intact in nearly all strains of pneumococcus, 75% of sequenced pneumococcal strains contain mutations that disrupt the *blpA* and/or *blpB* genes (14). Strains harboring these mutations are predicted to be incapable of making functional BlpAB transporter. These BlpAB$^-$ strains are dependent on *com* activation to activate their own *blp* system (*com*-dependent *blp* activation) through the *com* to *blp* crosstalk described above (39, 40). They can also respond to BlpC pheromone secreted by other strains since their BlpHR sensory apparatus remains intact. However, it is not clear if BlpAB$^-$ strains can effect pneumocin-mediated killing. Pneumocins, like the pheromones CSP and BlpC, need to be processed and secreted before they can exert their effects. BlpAB is the putative transporter for pneumocins, so loss of BlpAB would be expected to render cells incapable of secreting

active pneumocins. Two previous studies have reported conflicting results on whether BlpAB$^-$ strains exhibit pneumocin-mediated inhibitory activity (14, 137).

At the heart of both the *com* and *blp* systems lie the peptide transporters ComAB and BlpAB, which belong to the superfamily of ABC (ATP-binding cassette) transporters. ComAB and BlpAB are formed from A and B components. The B components, encoded by *comB* and *blpB*, serve unknown functions. Meanwhile, the A components, encoded by *comA* and *blpA*, contain all the known functional domains. Each A component contains an N-terminal peptidase domain (PEP) followed by the channel-forming transmembrane domain (TMD) and a C-terminal nucleotide-binding domain (NBD) (4). Two copies of the A component are required to form a functional transporter unit. The presence of PEP places ComAB and BlpAB into a family of ABC transporters called peptidase-containing ABC transporters (PCATs). PEP catalyzes the first step of transport called *processing*: the cleavage of the peptide substrate immediately C-terminal to a conserved double-glycine (GG) motif (sometimes also GA or GS), thereby removing the substrate's N-terminal signal sequence (4, 8). After processing, the now-mature peptide is secreted out of the cell using energy provided by the ATP-hydrolyzing NBD. While ComAB's ability to secrete both ComC and BlpC has revealed some degree of promiscuity in its substrate selection process, it is unknown if BlpAB shares this property or just how far it extends.

In this study, we demonstrate that BlpAB can secrete CSP and show the functional significance of this in the 25% of strains that are BlpAB$^+$. In addition, we demonstrate that ComAB can process and secrete pneumocins, suggesting that BlpAB$^-$ strains can support pneumocin-mediated inhibition, but only under competence-permissive conditions. We also show that temporal regulation of ComAB and BlpAB expression lead to differences in pneumocin transport efficiency in BlpAB$^+$ and BlpAB$^-$ strains. Finally, we demonstrate that BlpAB$^+$ strains are more effective competitors than BlpAB$^-$ strains in a mouse nasopharyngeal colonization model. These findings suggest that the diversification of the pneumococcal population into BlpAB$^+$ and BlpAB$^-$ types occurred to allow for a small pool of costly antagonistic strains that can eliminate competitors when environmental conditions mandate this approach.

2.3 Results

2.3.1 Development of a com/blp dual luciferase reporter.

To facilitate the study of *com-blp* crosstalk, we developed a dual luciferase reporter system to monitor *com* and *blp* activation in the same population of cells in real time. For our two luciferases, we chose the NanoLuc luciferase (Nluc, Promega) and the red-emitting *Luciola italica* luciferase (RFluc, Targeting Systems). We selected these two candidates because their emission spectra have little overlap and can be conveniently separated using optical filters (Fig. 2.2A), and they use different substrates with no cross-reactivity (181, 182). We also validated that the Nluc substrate, furimazine

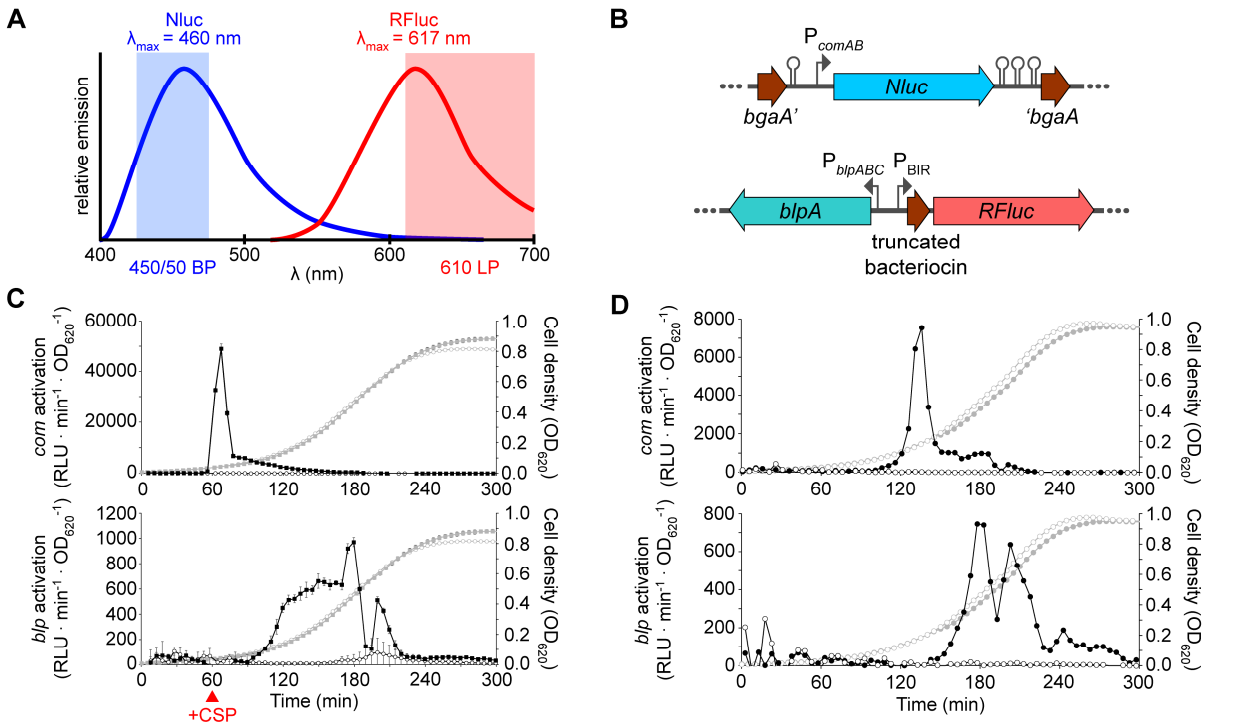

Figure 2.2. Dual *com-/blp-* luciferase reporter allows simultaneous monitoring of com and blp activation kinetics. (A) Diagram of the emission spectra of Nluc (blue) and RFluc (red) and the optical filter sets used to separate their signals. BP: band-pass; LP: long-pass. **(B)** Schematic representation of P_{comAB}-*Nluc* and P_{BIR}-*RFLuc* constructs. Bent arrows denote promoters. Hairpins denote transcriptional terminators. **(C)** An R6-derived wild-type dual reporter strain was grown in a 96-well plate in THY pH 7.1. Cells were monitored for growth (right y axes, light shading), *com* activation (top, left y axis, dark shading), and *blp* activation (bottom, left y axis, dark shading). At t = 60 min, cells were treated with either mock treatment (open circles) or 100 ng/mL CSP (closed squares). Data are plotted as the average ± S.D. of 4 wells. **(D)** An R6-derived wild-type dual reporter strain was grown in a 96-well plate in THY pH 7.4 without treatment. Cells were monitored as in panel C. Data from two representative wells are shown, one with spontaneous *com* and *blp* activation (filled circles), and one without (open circles).

(a.k.a. Nano-Glo substrate) was suitable for use in live pneumococcal cultures (Fig. 2.3).

We created a transcriptional fusion of Nluc to the ComE-regulated *comA* promoter (P_{comA}) and a transcriptional fusion of RFluc to the BlpR-regulated proximal BIR promoter (P_{BIR}) (Fig. 2.2B). These two constructs were then transformed into "wild-type" unencapsulated (R6) and encapsulated (D39) strains engineered to encode intact versions of *blpAB*. In these dual reporters Nluc activity correlates with *com* activation and RFluc activity correlates with *blp* activation.

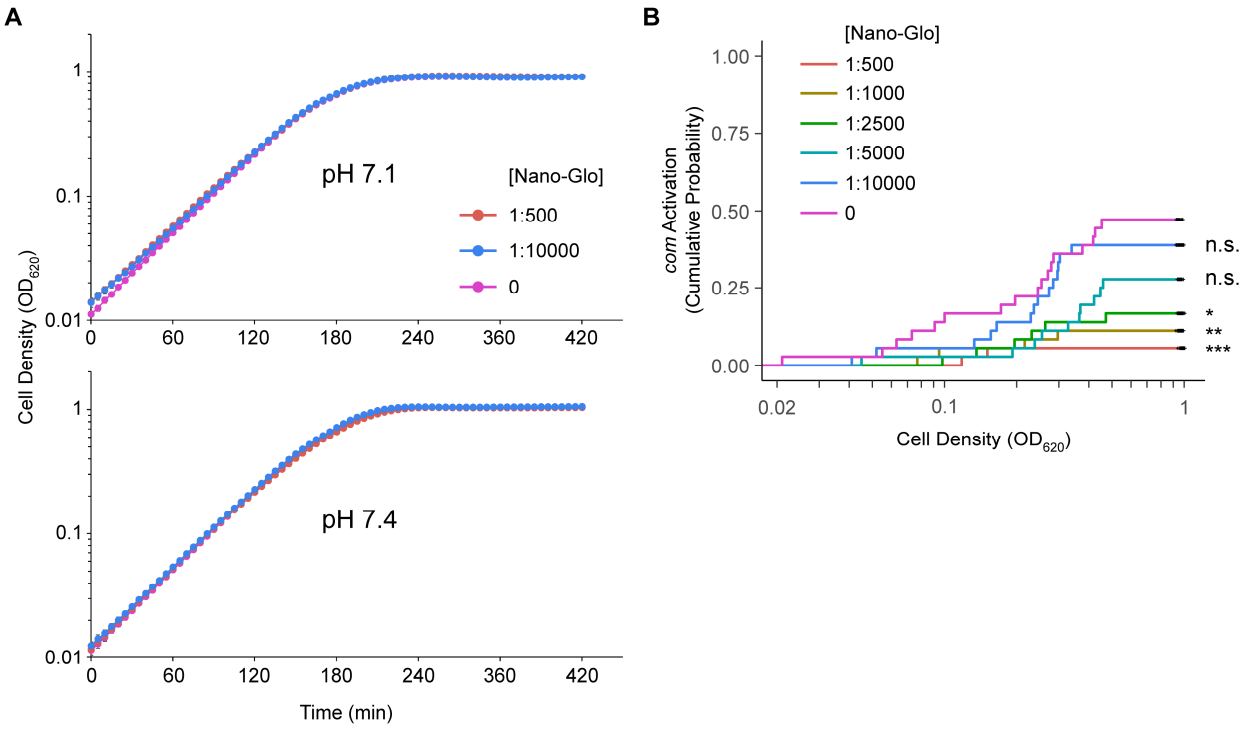

Figure 2.3. Validation of Nano-Glo substrate (furimazine). (A) Effect of Nano-Glo substrate on cell growth. An R6-derived Δ*comC* strain was grown in a 96-well plate in THY pH 7.1 (A) and pH 7.4 (B) containing 165 μM D-luciferin and various dilutions of Nano-Glo substrate (furimazine). Growth was monitored continuously. A Δ*comC* strain was used to prevent spontaneous competence activation, which can confound results by altering growth kinetics independently of Nano-Glo concentration. Data are plotted as the average ± S.D. of 12 wells. **(B)** Effect of Nano-Glo substrate on spontaneous *com* activation. An R6-derived P_{comAB}-*luc* reporter strain was grown in a 96-well plate in THY pH 7.4 containing 165 μM D-luciferin and various dilutions of Nano-Glo substrate (furimazine). Each well was continuously monitored for growth and *com* activation. Data for *com* activation events as a function of cell density were fit to the Kaplan-Meier estimator. Wells that did not experience a *com* activation event before cells reached their maximum density were censored (crosses). N=36 wells per Nano-Glo dilution, pooled from 3 independent experiments. Statistics: *** $p < 0.001$, ** $p < 0.01$, ** $p < 0.05$, all comparisons vs. [Nano-Glo] = 0, log-rank test with Holm correction for multiple comparisons.

We confirmed that the dual reporter successfully recapitulates what has been previously published regarding the response of the *com* and *blp* systems to CSP treatment (39, 40, 88, 116) (Fig. 2.2C). Using it, we are also able to detect spontaneous *com* and *blp* activation events when cells are grown under permissive conditions (Fig. 2.2D). In conclusion, the dual reporter provides accurate and convenient simultaneous readouts of *com* and *blp* activation in live cells.

2.3.2 BlpAB processes and secretes ComC/CSP.

We used the dual luciferase reporter to assay the response of cells to treatment with synthetic BlpC pheromone during growth in THY broth at pH 7.1, which is non-permissive for spontaneous *com* activation. We noticed that BlpC treatment led to activation of the *com* system (Fig. 2.4A). This BlpC-induced *com* activation was delayed by roughly 30 minutes compared to activation of the *blp* system, which responds almost immediately to BlpC. Given this delay, we reasoned that it was unlikely this *com* activation resulted from direct upregulation of P_{comA} by BlpR. Moreover, BlpC-induced *com* activation is completely abolished in the Δ*blpA* and Δ*comC* mutants but persists in the Δ*comAB* and Δ*blpC* mutants (Fig. 2.4A and B). These data led us to hypothesize that the mechanism underlying the cross-activation was BlpAB-mediated CSP secretion.

To test our hypothesis, we employed a previously developed peptide processing assay (183). Using FLAG-tagged peptides, the unprocessed and processed forms can be separated by size using SDS-PAGE and detected via western blot. FLAG-tagged peptides cannot be secreted by ComAB/BlpAB and are retained within the cytoplasm (183), most likely due to the high charge density of the FLAG tag interfering with loading into the transporter channel. Therefore, this assay only assesses the processing step of transport.

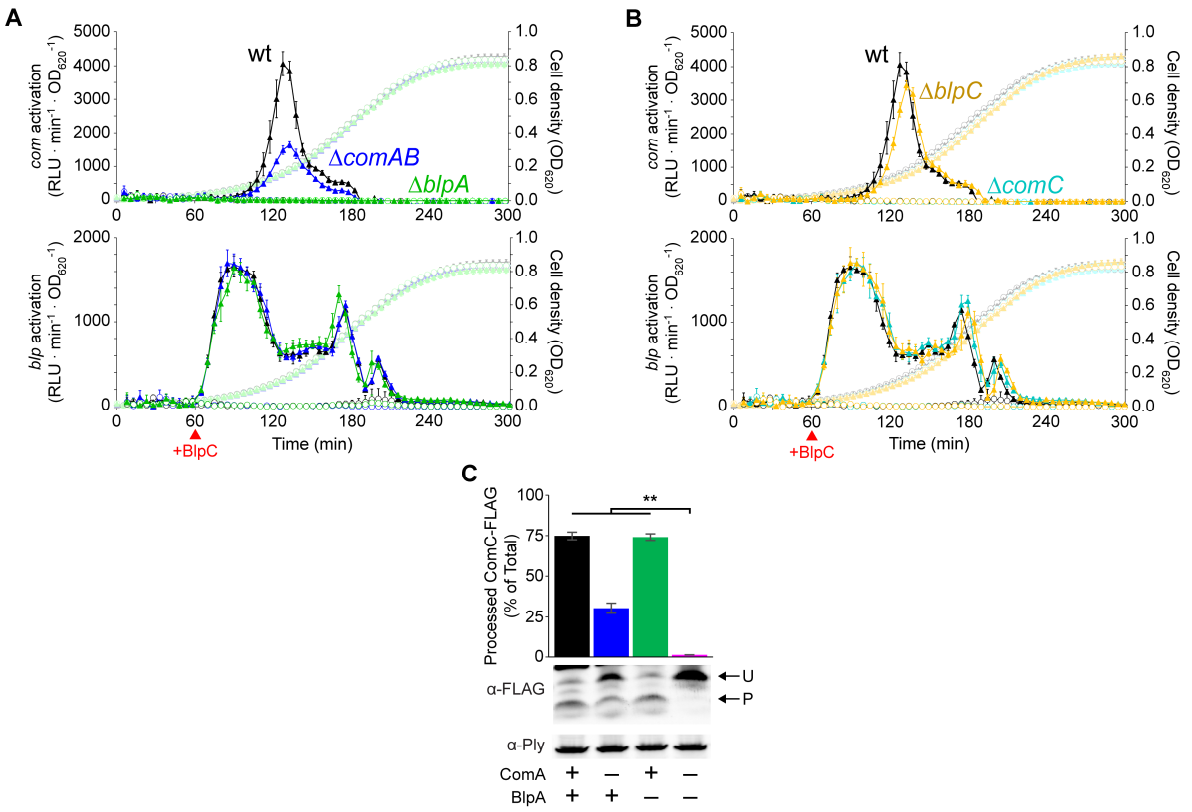

Figure 2.4. BlpAB processes and secretes ComC/CSP. (A, B) R6-derived dual reporter strains were grown in 96-well plates in THY pH 7.1. Cells were monitored for growth (right y axes, light shading), *com* activation (top, left y axis, dark shading), and *blp* activation (bottom, left y axis, dark shading). At t=60 min, cells were treated with either mock treatment (open circles) or 100 ng/mL BlpC (closed triangles). Data from one representative experiment are shown and are plotted as the average ± S.D. of 4 wells. **(C)** ComC-FLAG processing in wt (black), Δ*comAB* (blue), Δ*blpA* (green), and Δ*comAB*Δ*blpA* (magenta) strains. Cells were grown in THY pH 7.1, induced with 100 ng/mL CSP and 100 ng/mL BlpC, and whole-cell lysates were collected 30 min later for western blot. Representative blots are presented, showing the unprocessed (U) and processed (P) forms of ComC-FLAG and a pneumolysin (Ply) loading control. The amount of processed ComC-FLAG as a percentage of total (processed and unprocessed) ComC-FLAG is quantified and plotted as the average ± S.E. of 4 independent experiments. Statistics: ** $p < 0.01$, ANOVA with Tukey HSD.

We engineered R6 strains expressing C-terminally FLAG-tagged ComC (ComC-FLAG) in place of wild-type ComC from the native *comC* locus. Using these strains, we assayed for ComC processing in the wild-type, Δ*comAB*, Δ*blpA*, and Δ*comAB*Δ*blpA* backgrounds 30 minutes after induction with CSP and BlpC in THY pH 7.1 (Fig. 2.4C). The double mutant was included as a control to evaluate for the presence of a transporter or transporters other than ComAB or BlpAB that might contribute to ComC processing. The double mutant showed only a negligible amount of ComC processing, suggesting there are no other such transporters. Meanwhile, we

confirmed that ComAB can process ComC, consistent with established models of *com* regulation. We also observed ComC processing in the Δ*comAB* strain expressing only BlpAB. These results indicate that BlpAB can at least complete the first step of ComC transport (processing) and when taken together with the *com* activation kinetics data (Fig. 2.4A and B) strongly suggest that BlpAB can carry out the second step (secretion) as well. We conclude that BlpAB can transport ComC/CSP and this mechanism is responsible for the observed BlpC-induced *com* activation.

2.3.3 ComAB and BlpAB both process and secrete pneumocins.

Having established that ComAB and BlpAB both secrete CSP and BlpC (39, 40), we sought to determine if the promiscuity of these transporters extended to other substrates as well. The current model of substrate recognition by ComAB/BlpAB posits that the peptidase domain of the transporters interacts with the substrates' N-terminal signal sequences. In addition to the double-glycine motif, four specific hydrophobic residues in the signal sequence (Fig. 2.5A, yellow highlights) are important for this interaction (8). These residues are conserved across all *com*- and *blp*-regulated double-glycine peptides found in pneumococcus. Moreover, the residues in the transporter peptidase domain that are thought to participate in substrate recognition (5) are also highly conserved in both ComA and BlpA (Fig. 2.5B, yellow highlights). Given this, we hypothesized that ComAB and BlpAB recognize and transport the same pool of substrates, including the pheromones ComC/CSP and BlpC, the pneumocins, and the competence-induced bacteriocins CibAB.

We were particularly interested to test this hypothesis on the pneumocins for several reasons. First, while long suspected to be BlpAB, the transporter or transporters responsible for pneumocin secretion have never been definitively identified. Second, the question of whether ComAB can secrete the pneumocins is an important unanswered question given the ability of the *com* system to upregulate pneumocin expression and the absence of BlpAB in 75% of strains. Since all pneumocins share very similar signal sequences (Fig. 2.5A), we reasoned that the details of transport would be similar, if not identical, across the different pneumocins. As such, to use for subsequent experiments

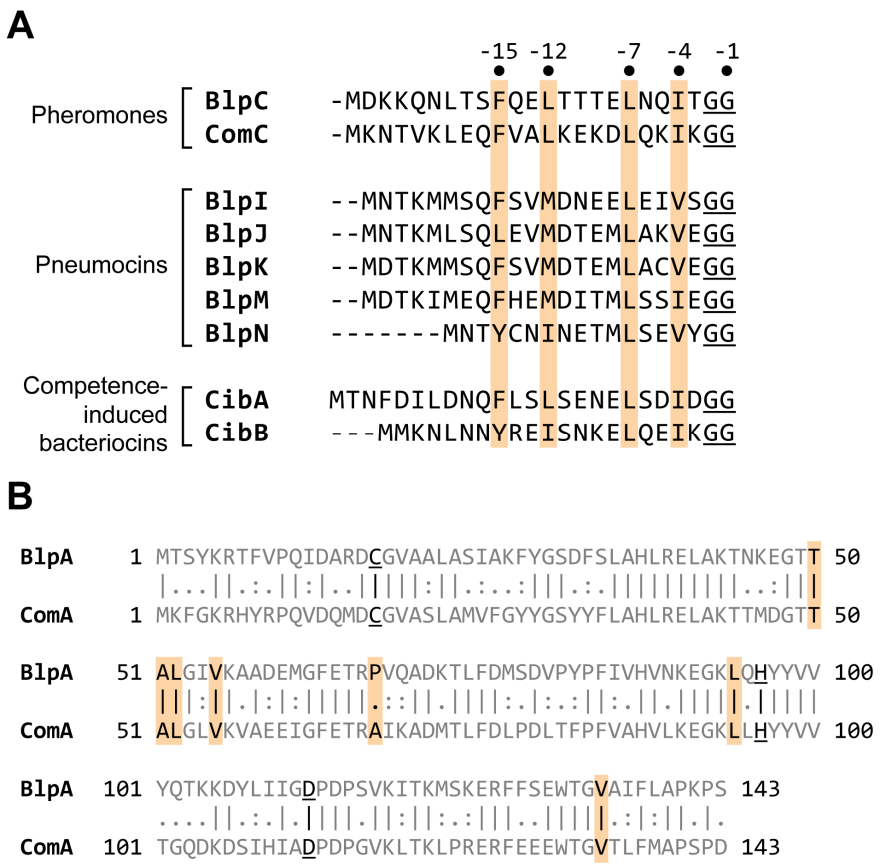

Figure 2.5. Pneumococcal *com*- and *blp*-regulated GG peptides and their transporters share conserved motifs important for substrate recognition. (A) Alignment of the N-terminal signal sequences of selected competence- or *blp*-related double-glycine peptides found in *S. pneumoniae*. The double-glycine motif is underlined. Residues predicted to interact with the transporter peptidase domain are highlighted in yellow. **(B)** Alignment of the transporter peptidase domains of ComA and BlpA from *S. pneumoniae*. The catalytic triad is underlined. Residues predicted to interact with substrates are highlighted in yellow.

we chose a representative pneumocin, $BlpI_{P133}$ (from now on referred to as BlpI), which has been previously shown to inhibit sensitive strains when expressed with its partner, BlpJ (14).

To assess for BlpI secretion, we chose to employ the HiBiT tag detection system (Promega). The HiBiT tag is a short, 11-residue peptide tag (VSGWRLFKKIS) that associates with the inactive LgBiT luciferase fragment with sub-nanomolar affinity to complement the latter's luciferase activity (184, 185). This system allows for highly sensitive detection of HiBiT-tagged peptides and proteins using a bioluminescence assay (185).

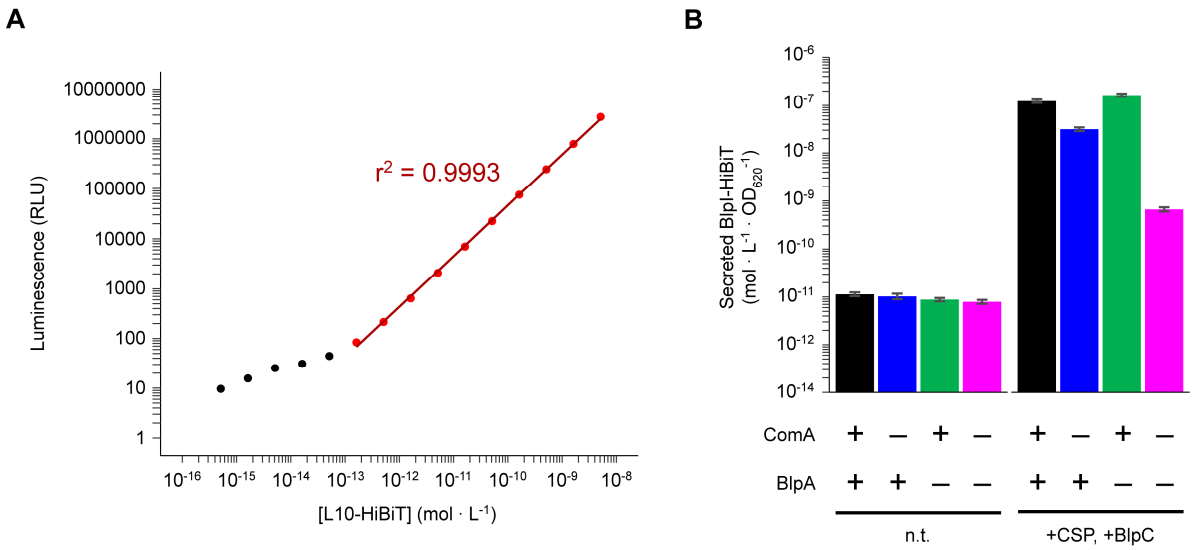

Figure 2.6. Validation of BlpI-HiBiT secretion assay. (A) Dose-response plot of HiBiT detection assay using synthetic L10-HiBiT peptide diluted in sterile THY pH 7.1. Dark red line represents linear regression model created using data points highlighted in red. The assay exhibits a linear response over 4+ orders of magnitude (from ~100 fM to 5 nM HiBiT). Quantification of HiBiT concentrations greater than 5 nM is primarily limited by the fact that the complemented luciferase consumes the furimazine substrate too quickly, leading to rapid signal degradation. **(B)** BlpI-HiBiT secretion in uninduced and induced cells. BlpI-HiBiT expressing strains were induced with either mock-treatment (left) or 200 ng/mL CSP and 200 ng/mL BlpC (right). One hour after induction, supernatants were collected and assayed for BlpI-HiBiT. Data are plotted as the average ± S.E. of 3 independent experiments.

After validating the assay system (Fig. 2.6), we used R6 strains expressing BlpI-HiBiT from the native, BlpR-regulated proximal BIR promoter to evaluate BlpI secretion 60 minutes after BlpC and CSP induction in the wild-type and transporter deletion backgrounds (Fig. 2.7A). The amount of BlpI-HiBiT detected in the supernatant of the double mutant was 180-fold less than that of the wild-type, confirming that ComAB and BlpAB are the primary contributors to BlpI secretion. We also found that both single mutants secreted significantly more BlpI-HiBiT than the double mutant. In contrast, we detected similar levels of BlpI in the cell-associated fractions of all strains (Fig. 2.7A, bottom panel). We corroborated these findings using a BlpI-FLAG processing assay (Fig. 2.7B), which showed high levels of processing in both single transporter mutants. Therefore, both ComAB and BlpAB can process and secrete BlpI and likely the rest of the pneumocins as well.

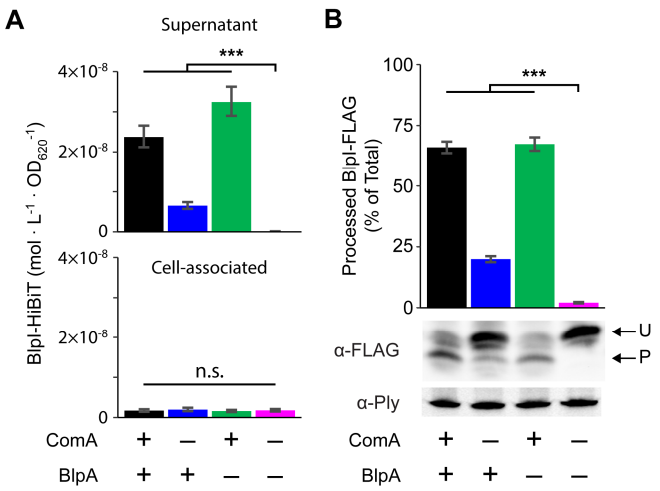

Figure 2.7. ComAB and BlpAB process and secrete the pneumocin BlpI. (A) BlpI-HiBiT secretion in wt (black), Δ*comAB* (blue), Δ*blpA* (green), and Δ*comAB*Δ*blpA* (magenta) strains. Concentration of BlpI-HiBiT normalized to cell density detected in the supernatant (top) and cell-associated (bottom) fractions 60 minutes after induction with 200 ng/mL CSP and 200 ng/mL BlpC in THY pH 7.1. Plotted as the average ± S.E. of 3 independent experiments. **(B)** BlpI-FLAG processing. Cells were grown in THY pH 7.1, induced with 100 ng/mL CSP and 100 ng/mL BlpC, and whole-cell lysates were collected 30 minutes later. Representative blots are presented, showing the unprocessed (U) and processed (P) forms of BlpI-FLAG and a pneumolysin (Ply) loading control. The amount of processed BlpI-FLAG as a percentage of total (processed and unprocessed) BlpI-FLAG is quantified and plotted as the average ± S.E. of 4 independent experiments. Statistics: n.s., not significant; *** $p < 0.001$; ANOVA with Tukey HSD.

2.3.4 Pheromone secretion by BlpAB enhances competence activation and allows competence-independent blp activation.

Having determined that ComAB and BlpAB are functionally redundant for secretion of pheromones and pneumocins, we sought to determine what advantages, if any, intact BlpAB confers to pneumococcus. First, we hypothesized that ComC/CSP secretion by BlpAB could drive spontaneous *com* activation, similar to how BlpC secretion by ComAB could drive spontaneous *blp* activation; if true, then we should be able to observe spontaneous *com* activation following *blp* activation in a Δ*comAB* mutant. To test this, we used our dual reporter to monitor spontaneous *com* activation in a panel of transporter and pheromone deletion mutants grown in THY pH 7.4 in a 96-well plate. Each well was independently assessed for *com* activation events, defined as when the *com* activation level of the well crossed above an empirically determined threshold value. If a well reached its maximum observed cell density as measured by OD_{620} before a *com* activation event occurred, then it was censored. Finally, the event

data were fit to Kaplan-Meier estimators and plotted as the cumulative probability of *com* activation versus cell density.

Under these conditions, we observed within-strain stochasticity in the pattern of *com* activation; the unencapsulated R6 wild-type strain activated *com* in only a subset of identically inoculated wells, and there was large variation in the timing of *com* activation among those wells in which it did occur (Fig. 2.8A). Contrary to our hypothesis, we did not observe *com* activation in the Δ*comAB* mutant. However, we did observe a defect in *com* activation in the Δ*blpA* mutant compared to the wild-type strain. This defect was not caused by an upstream deficiency in *blp* activation, since *blp* activation was not observed prior to *com* activation in any strain. Due to conflicting reports on the effect of capsule on CSP signaling (103, 104), we tested whether we could reproduce this

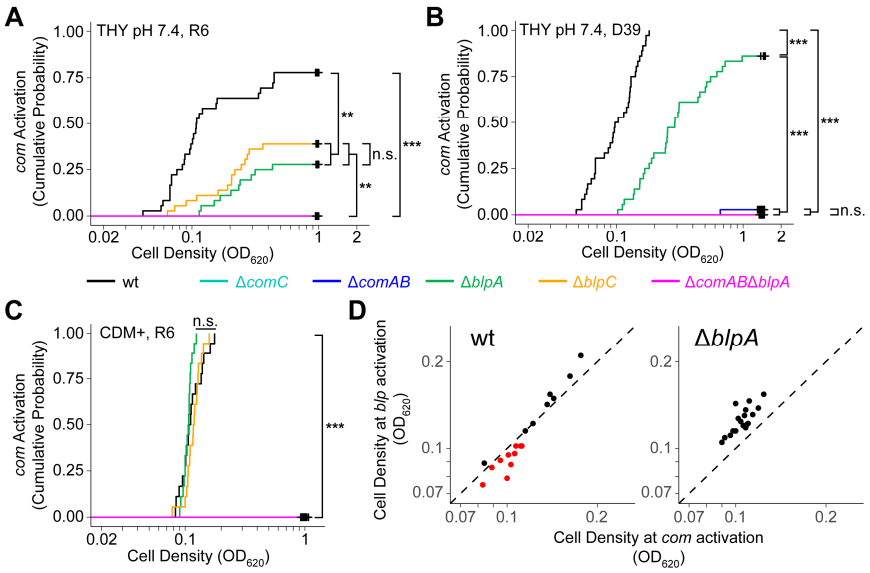

Figure 2.8. BlpAB enhances spontaneous *com* activation and drives *com*-independent *blp* activation. (A-C) Spontaneous *com* activation in THY and CDM+. R6 (A,C) and D39 (B) dual reporter strains were grown in 96-well plates in THY pH 7.4 (A,B) or CDM+ pH 7.0 (C). All six (wild-type, pheromone and transporter deletion) strains are represented in (A,C); only the wild-type and transporter deletion strains are represented in (B). Each well was continuously monitored for growth and *com* activation. Data were fit to the Kaplan-Meier estimator. Wells that did not experience a *com* activation event before cells reached their maximum density were censored (crosses). N = 36 (A,B) or 18 (C) wells per strain, pooled from 3 independent experiments. Statistics: *** $p < 0.001$, ** $p < 0.01$, n.s. not significant; log-rank test with Holm correction for multiple comparisons. **(D)** Cell density at *com* activation plotted against cell density at *blp* activation in individual wells from panel C of the wt and Δ*blpA* strains. Points that fall on the dotted line represent wells in which *blp* activation occurred at the same time as *com* activation. Points that fall to the left of the dotted line represent *blp* activation occurring after *com* activation. Points that fall to the right of the dotted line (highlighted in red) represent *blp* activation occurring before *com* activation.

phenotype in an encapsulated strain. Indeed, we observed the same competence defect in the Δ*blpA* mutant compared to the wild-type strain in the encapsulated D39 background (Fig. 2.8B). Therefore, under these conditions the presence of capsule does not appreciably affect BlpAB-mediated *blp* to *com* crosstalk.

To assess the contribution of *com* and *blp* activation states of the starter cultures used to inoculate media at the beginning of the assay, we repeated the previous experiment using inocula from starter cultures of R6 strains grown in THY pH 6.8, a condition under which neither *com* nor *blp* activation occurs. This produced results that closely resemble those obtained using pH 7.4 starter cultures (Fig. 2.9).

We next tested whether enhanced *com* activation by BlpAB occurred under other growth conditions. We repeated the spontaneous *com* activation experiment using CDM+, a more minimal medium compared to THY that is used for maintaining pneumococcal biofilms (105). Unlike in THY, we did not observe a *com* activation defect in the R6 Δ*blpA* mutant in CDM+ compared to the wild-type (Fig. 2.8C). From the same experiment in CDM+, we were also able to gather data on *blp* activation using the dual reporter. When analyzing these data, we noticed that a significant number of wells of the wild-type strain activated *blp* before *com* (Fig. 2.8D, left panel, red dots). This *com*-

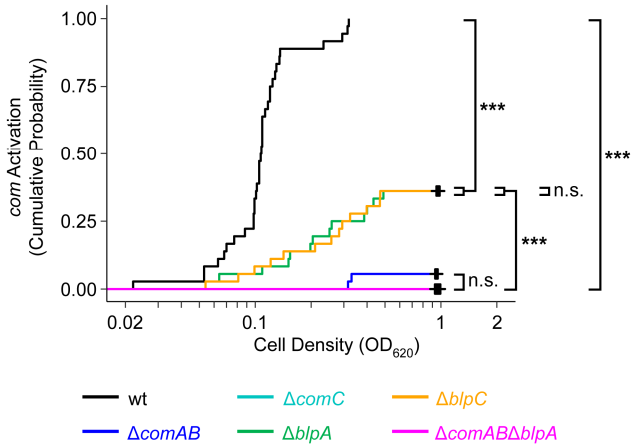

Figure 2.9. Spontaneous *com* activation in R6 dual reporters inoculated from pH 6.8 starter cultures. R6 dual reporter strains inoculated from starter cultures grown in THY pH 6.8 were grown in 96-well plates in THY pH 7.4. Each well was continuously monitored for growth and *com* activation. Data were fit to the Kaplan-Meier estimator. Wells that did not experience a *com* activation event before cells reached their maximum density were censored (crosses). N = 36 wells per strain, pooled from 3 independent experiments. Statistics: *** $p < 0.001$, n.s. not significant; log-rank test with Holm correction for multiple comparisons.

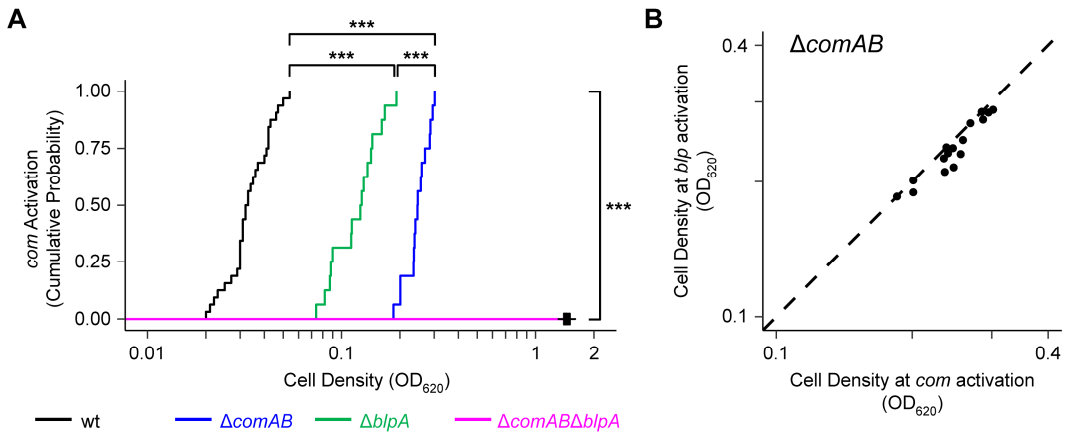

Figure 2.10. BlpAB enhances *com* activation and drives *blp*-dependent *com* activation in the D39 strain during growth in CDM+. (A) D39 dual reporter strains inoculated from starter cultures grown in THY pH 6.8 were grown in 96-well plates in CDM+ pH 7.0. Each well was continuously monitored for growth and *com* activation. Data were fit to the Kaplan-Meier estimator. Wells that did not experience a *com* activation event before cells reached their maximum density were censored (crosses). N = 32 (wt) or 16 (all others) wells per strain, pooled from 4 independent experiments. Statistics: *** $p < 0.001$, n.s. not significant; log-rank test with Holm correction for multiple comparisons. **(B)** Cell density at *com* activation plotted against cell density at *blp* activation in individual wells from panel A of the Δ*comAB* strain. Points that fall on the dotted line represent wells in which *blp* activation occurred at the same time as *com* activation. Points that fall to the right of the dotted line represent *blp* activation occurring before *com* activation.

independent *blp* activation occurred in just over half of the wells (10/18) in the wild-type strain compared to zero wells in the Δ*blpA* mutant (Fig. 2.8D, right panel). Consistent with previous work (40), these results show that BlpAB can promote and is required for *com*-independent *blp* activation.

We repeated the CDM+ experiment with the D39 strain and found that in this strain background BlpAB does enhance *com* activation during growth in CDM+ (Fig. 2.10A). In addition, we observed *com* activation in all the wells containing the Δ*comAB* mutant (Fig. 2.10A, blue). These *com* activation events were coincident with or immediately followed *blp* activation (Fig. 2.10B). Therefore, under these conditions, BlpAB in the D39 strain can drive *com* activation independently of ComAB during periods of *blp* activation. This mirrors how ComAB can drive *blp* activation in BlpAB$^-$ strains during *com* activation.

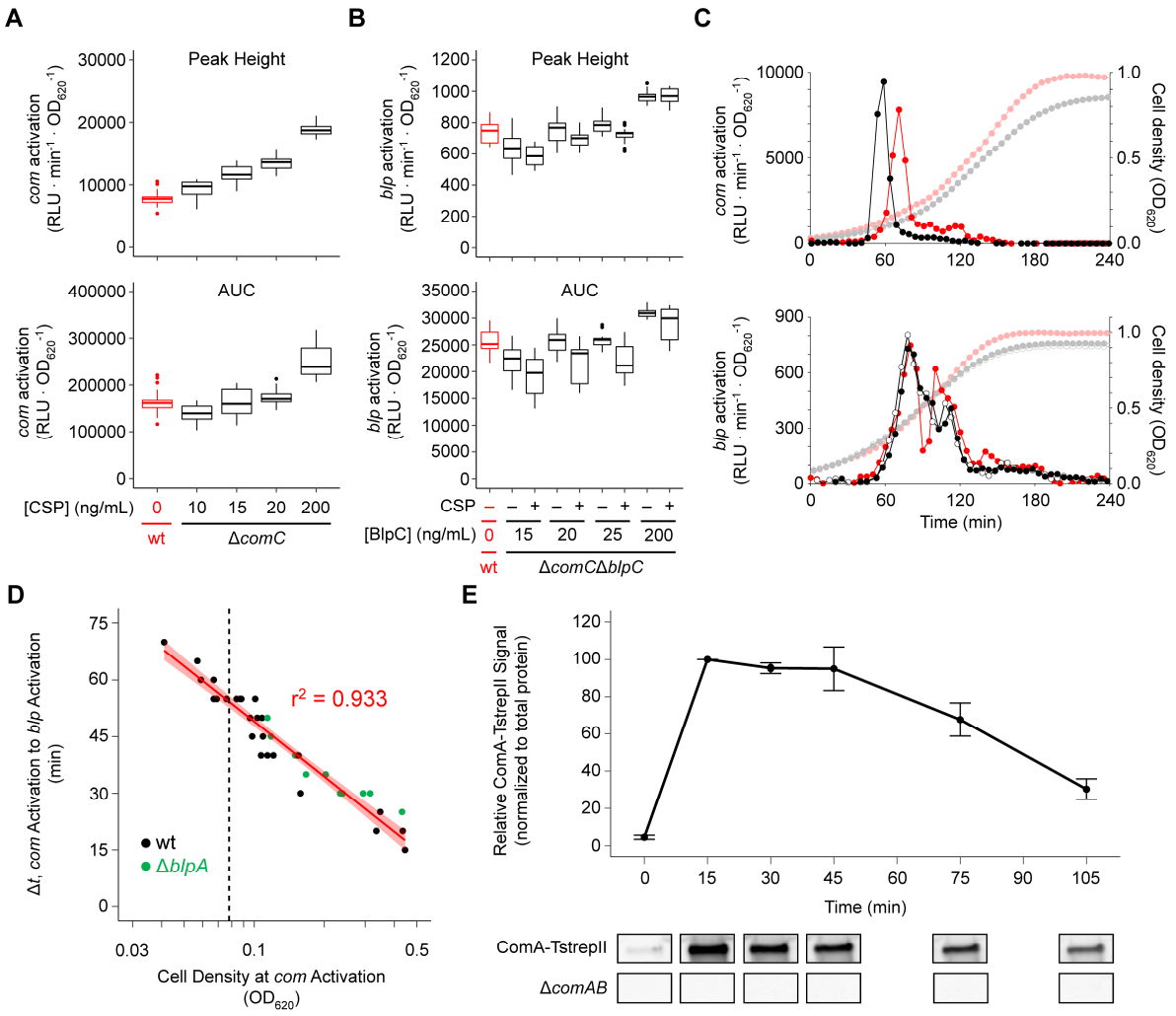

Figure 2.11. Parameters of natural *com* and *blp* activation. (A) *com* activation characteristics. An R6 Δ*comC* strain was induced with various concentrations of CSP at OD_{620} 0.078 (black, N = 24 per condition, pooled from 2 independent experiments) and the peak height and area under the curve (AUC) of *com* activation from 0 to 105 minutes post-induction is plotted here. For comparison, data for spontaneous *com* activation of an R6 wild-type strain in THY pH 7.4 (red, N = 17), taken from the same experiments as those in Fig. 5A and restricted to *com* activation events occurring between OD_{620} 0.05 and 0.1, are also included. **(B)** *blp* activation characteristics. An R6 Δ*comC*Δ*blpC* strain was induced with either 0 or 10 ng/mL CSP at OD_{620} 0.078 and then various concentrations of BlpC 45 minutes later (black, N = 18 per condition, pooled from 3 independent experiments) and the peak height and area under the curve (AUC) of *blp* activation from 0 to 105 minutes post-CSP induction is plotted here. For comparison, data for spontaneous *blp* activation of an R6 wild-type strain in THY pH 7.4 (red, N = 17), collected from the same wells as those in panel A, are also included. **(C)** Representative examples of *com* (top) and *blp* (bottom) activation kinetics from panels A and B, respectively. Spontaneous activation of the wild-type strain is shown in red, and CSP/BlpC-induced activation of the pheromone deletion strains is shown in black (open circles: -CSP; closed circles: +CSP). *com*/*blp* activation levels are shown in dark shading (left y-axis) and cell density is shown in light shading (right y-axis). The timelines of the strains in each graph were aligned so that *com* activation begins to occur at roughly the same timepoint for all strains. **(D)** Correlation of the time delay between spontaneous *com* and *blp* activation with cell density at the point of *com* activation in wild-type (black) and Δ*blpA* (green) strains in THY pH 7.4. Data were taken from the same experiments as in Fig. 2.8A and represent all wells in which *blp* activation was observed. Red line and shading represent linear regression estimate ± 95% CI. The dashed vertical line is placed at OD_{620} 0.078. **(E)** Twin-Strep-tagged ComA (ComA-TstrepII) protein levels following treatment

with 10 ng/mL CSP at OD_{620} 0.078 in THY pH 7.1. Samples were collected at the indicated timepoints, and membrane fractions were assayed by quantitative western blot with anti-Strep-tag II antibody. Blots of samples from one representative experiment of the strain expressing ComA-TstrepII (top) and of a Δ*comAB* negative control (bottom) are shown. For quantification, ComA signal was normalized to total protein and expressed relative to the 15-minute time point, which was set to a value of 100. Data are plotted as the average ± S.E. of 3 independent experiments.

2.3.5 BlpAB ensures efficient transport of pneumocins.

Lastly, we wanted to investigate whether $BlpAB^+$ strains, which possess both ComAB and BlpAB, enjoy advantages in pneumocin secretion over $BlpAB^-$ strains, which possess only ComAB. The previous BlpI secretion assay (Fig. 2.7) was performed on cells simultaneously induced with saturating concentrations of both CSP and BlpC. These conditions are unlikely to resemble what occurs during natural, spontaneous activation of *com* and *blp*.

First, we found that the pheromone concentrations required to induce strains to levels and with kinetics similar to spontaneous activation were 10 ng/mL for CSP and 25 ng/mL for BlpC (Fig. 2.11A-C), much lower than the 100-200 ng/mL used in previous processing and secretion assays. Second, in the absence of an exogenous source of BlpC, $BlpAB^-$ strains can only activate *blp* in a *com*-dependent fashion. This *com*-dependent *blp* activation also happens in $BlpAB^+$ strains. In both strain backgrounds, *blp* activation in this manner occurs after *com* activation following a cell density-dependent delay (Fig. 2.11D), the length of which is not affected by the presence of BlpAB ($p > 0.05$, ANCOVA). Third, we observed that ComA protein levels following CSP induction rapidly decrease after 45 minutes (Fig. 2.11E). This is consistent with the rapid shut-off in transcription activity from P_{comA} seen with the luciferase reporter (Fig. 2.11C). Given these data, we hypothesized that $BlpAB^-$ strains, which must rely on ComAB for secretion, secrete less pneumocins during *com*-dependent *blp* activation than $BlpAB^+$ strains, which can use both ComAB and BlpAB.

To test this hypothesis, we assayed BlpI secretion under conditions similar to spontaneous *com* and *blp* activation. The strains used in this experiment were engineered to express pheromone-receptor mismatched CSP2-ComD1 and $BlpC_{6A}$-$BlpH_{R6}$ pairs. Therefore, these strains are deficient in auto-activation of both *com* and

blp to ensure differences in transporter content (and hence pheromone secretion) did not affect activation kinetics. Cells were induced with either BlpC alone (to mimic *com*-independent *blp* activation), CSP and BlpC together ("simultaneous"), or CSP followed by BlpC 45 minutes later ("staggered", to mimic *com*-dependent *blp* activation) (Fig. 2.12A). Consistent with our hypothesis, we observed a small but statistically significant increase in BlpI secretion from the wild-type strain compared with the Δ*blpA* mutant in the staggered treatment group (Fig. 2.12B). This difference was not seen in the simultaneous treatment group. Additionally, when comparing the staggered to the simultaneous treatment group, a large decrease in BlpI secretion was only seen in the strains possessing ComAB, and of these two strains the wild-type strain which possesses both ComAB and BlpAB suffered the smaller decrease.

The differences seen between the wild-type and Δ*blpA* strains in the previous experiment were small when assessed at just 105 minutes after CSP treatment. However, given our observation that ComAB levels continuously decrease past 45 minutes post-CSP treatment, the differences should increase over time. To test this hypothesis, we monitored BlpI-HiBiT secretion over time following the "BlpC only" or "staggered" treatments (Fig. 2.12C). Consistent with falling ComAB levels, beginning at 60 minutes post-BlpC treatment (105 minutes post-CSP treatment) the Δ*blpA* mutant in the "staggered" treatment group showed increasingly large defects in BlpI-HiBiT secretion over time compared to the wild-type strain in the same treatment group. The same was seen when comparing the Δ*blpA* mutant in the "staggered" treatment group to the wild-type strain in the "BlpC only" treatment group. Lastly, in the "BlpC only" treatment group, only BlpAB-containing strains secreted BlpI at levels higher than the double transporter mutant, and the wild-type strain did not secrete more than the Δ*comAB* mutant ($p \geq 0.05$ at all timepoints). These data show that BlpI secretion through ComAB (but not BlpAB) rapidly decreases with time during *com*-dependent *blp* activation and is negligible during *com*-independent *blp* activation.

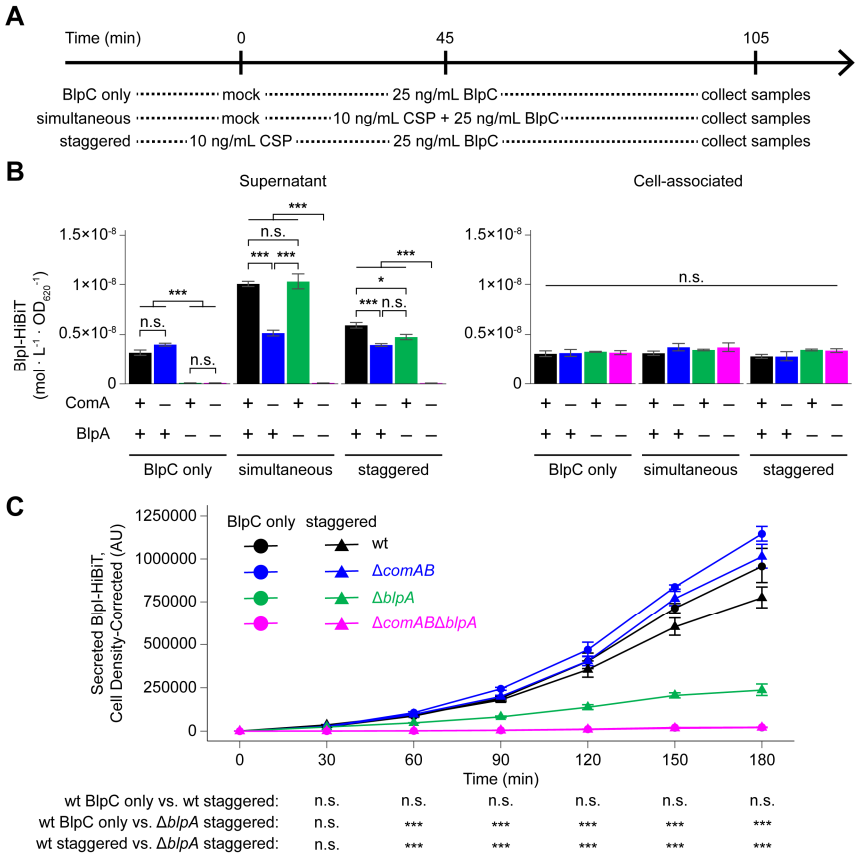

Figure 2.12. BlpAB promotes efficient pneumocin secretion. (A) Schematic of spontaneous activation-simulating BlpI-HiBiT secretion experiment. Treatment and sample collection regimen of the three different treatment groups are presented below the timeline. Time 0 was defined as when the cells reached OD_{620} 0.078. **(B)** BlpI secretion in wild-type (black), Δ*comAB* (blue), Δ*blpA* (green), and Δ*comAB*Δ*blpA* (magenta) strains grown in THY pH 7.1. Cells were treated as depicted in panel D and samples were collected and assayed for BlpI-HiBiT in the supernatant (left) and cell-associated (right) fractions. Data are plotted as the average ± S.E. of 3 independent experiments. Statistics: n.s. not significant; * $p < 0.05$, *** $p < 0.001$; ANOVA with Tukey HSD. **(C)** BlpI-HiBiT secretion over time in wt (black), Δ*comAB* (blue), Δ*blpA* (green), and Δ*comAB*Δ*blpA* (magenta) strains grown in THY pH 7.1. Beginning at OD_{620} 0.02 (t = -45 min), cells were given two treatments at -45 and 0 minutes: BlpC only, mock followed by 25 ng/mL BlpC; staggered, 10 ng/mL CSP followed by 25 ng/mL BlpC. Samples were collected every 30 minutes and assayed for BlpI-HiBiT in the supernatant fractions. Data were corrected for differences in cell density between samples independently per timepoint and are plotted as the average ± S.E. of 3 independent experiments. Green circles (Δ*blpA*, BlpC only) are obscured behind the magenta plots. Statistics: n.s. not significant; *** $p < 0.001$; ANOVA with Tukey HSD.

2.3.6 BlpAB⁺ strains have a competitive advantage over BlpAB⁻ strains during nasopharyngeal colonization.

We hypothesized that BlpAB⁺ strains' ability to activate *blp* independently of *com* and secrete greater amounts of pneumocins would give them a competitive advantage compared to BlpAB⁻ strains during nasopharyngeal colonization. To test this, we co-

inoculated mice with either a BlpAB$^+$ or BlpAB$^-$ pneumocin-expressing “killer” strain and a Δ*blp* pneumocin-sensitive strain and assessed competitive indices at 4 days post-inoculation. The BlpAB$^+$ killer strain outcompeted the sensitive strain to a greater extent than the BlpAB$^-$ killer strain did (Fig. 2.13A). We also examined how BlpAB$^+$ and BlpAB$^-$ strains fare in direct competition. We co-inoculated mice with pairs of either BlpC pherotype-matched or mismatched pneumocin-expressing BlpAB$^+$ and BlpAB$^-$ strains and assessed competitive indices at 4 days post-inoculation. The co-inoculated strains

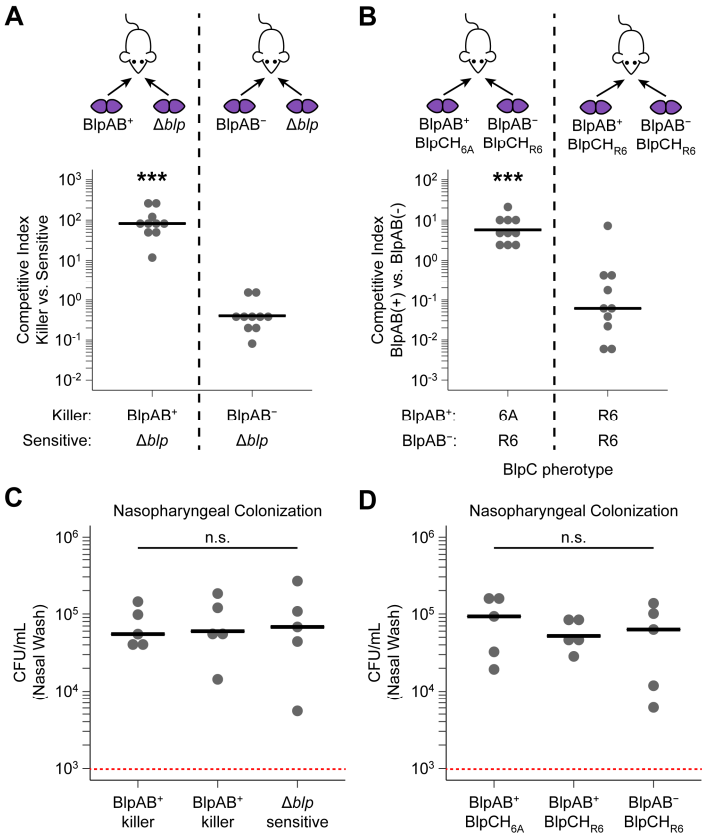

Figure 2.13. BlpAB$^+$ strains have a competitive advantage over BlpAB$^-$ strains during mouse nasopharyngeal colonization. All pneumococci are mouse-adapted serotype 19A strains. **(A)** Competition of BlpAB$^+$ and BlpAB$^-$ strains against a pneumocin-sensitive strain. BALB/c mice were co-inoculated in a 1:1 ratio with either a BlpAB$^+$ or BlpAB$^-$ pneumocin-expressing “killer” strain and a Δ*blp* (whole locus) sensitive strain. Both killer strains encode a P164-type BIR. CFU counts from the nasopharynx were obtained at 4 days post-inoculation and used to calculate competitive indices. N = 10 mice per competition. **(B)** Direct competition of BlpAB$^+$ strains against BlpAB$^-$ strains. Mice were co-inoculated as in panel A with either a BlpCH$_{6A}$ or BlpCH$_{R6}$-expressing BlpAB$^+$ strain and a BlpCH$_{R6}$-expressing BlpAB$^-$ strain. All strains encode a P133-type BIR. Competitive indices were calculated as in panel A. N=10 mice per competition. Black bars represent medians. Statistics: *** $p < 0.001$ (left-hand vs. right-hand group); Mann-Whitney test. **(C, D)** BALB/c mice were singly inoculated with mouse-adapted serotype 19A strains. CFU counts from nasal washes were obtained at 4 days post-inoculation. N = 5 mice per strain. Black bars represent medians. Statistics: n.s. not significant; Kruskal-Wallis test.

had identical BIRs; therefore, either strain could develop immunity to the other's pneumocins provided it activates *blp* at the appropriate time. We observed that the BlpAB$^+$ strain had a competitive advantage over the BlpAB$^-$ strain, but only when the two strains had mismatched BlpC pherotypes (Fig. 2.13B). All strains used in these competition assays colonize to similar levels when inoculated alone (Fig. 2.13C, D) indicating none have intrinsic colonization defects.

2.4 Discussion

We have presented evidence showing that competence and bacteriocin regulation in pneumococcus are more entwined than previously thought. While it was known that *com* could send positive inputs to the *blp* system, we show here that signals can also travel in the opposite direction due to secretion of CSP by BlpAB. While we found this *blp* to *com* crosstalk could drive *com* activation following or concurrently with *blp* activation (Fig. 2.10), the more common effect of the crosstalk was activation of *com* at higher frequencies and lower cell densities, even in the absence of overt *blp* activation (Fig. 2.8A, B). This highlights the importance of the basal level of transporter expression as the rate-limiting factor for CSP secretion and *com* activation; the addition of a second transporter (BlpAB) to augment basal CSP secretion by ComAB is sufficient to increase *com* activation. This phenotype was media and strain dependent; it was found in both the R6 and D39 strains during growth in THY but only in D39 during growth in CDM+. We also found that in THY, *com* activation suffered the same defect in the absence of BlpC signaling as in the absence of the BlpAB transporter (Fig. 2.8A; compare Δ*blpC* to Δ*blpA*). We speculate that basal levels of BlpC secretion and signaling lead to noise in *blp* activation during the pre-activation period. This noise results in transient elevated expression of BlpAB in a subpopulation of cells that leads to bursts of CSP secretion which over time help to push the total extracellular CSP concentration over the threshold required for *com* activation. Further studies will be needed to confirm this hypothesis and to further elucidate the determinants of transporter expression dynamics in the pre-activation period and their effects on *com* and *blp* activation.

We have also shown that the ability of both ComAB and BlpAB to secrete the other system's pheromone is not a special case, but rather a natural consequence of the transporters' promiscuity toward substrates. It was previously shown that purified peptidase domain from BlpA could cleave a synthetic CSP-like analog (186). We confirm here that full-length BlpAB can secrete native ComC/CSP in a live-cell context. We have also observed that both transporters can process and secrete the pneumocin BlpI (Fig. 2.7). These data support the conclusion that ComAB and BlpAB share the same substrate pool. Finally, our data indicate that the signal sequence-peptidase domain interaction is the primary factor that determines whether a peptide is secreted by ComAB/BlpAB. Both transporters tolerate a wide variety of mature peptides for secretion: small and amphipathic (CSP), charged (BlpC), and large and highly hydrophobic (pneumocins).

The promiscuity of ComAB and BlpAB have multiple functional consequences, which are lent added importance by the fact that 75% of pneumococcal strains lack a functional BlpAB but nearly all strains produce a functional ComAB. It was previously unclear whether BlpAB$^-$ strains could secrete pneumocins. We provide evidence here that these strains can in fact secrete pneumocins during *com*-dependent *blp* activation. Over short time frames, the amount of pneumocin secreted by BlpAB$^-$ strains in this manner is comparable to the amount secreted by BlpAB$^+$ strains during either *com*-independent or *com*-dependent *blp* activation (Fig. 2.12). This suggests that BlpAB$^-$ strains theoretically can effect pneumocin-mediated inhibition during *com*-dependent *blp* activation. This would be true for all but a small minority of strains that cannot activate *com* or produce ComAB due to acquired mutations in *com* regulatory genes (102). Obtaining direct evidence for this will be a priority for future studies.

Despite the above, BlpAB$^+$ strains still enjoy multiple advantages in pneumocin secretion over their BlpAB$^-$ counterparts due to differences in the regulation and kinetics of ComAB and BlpAB expression. First, BlpAB$^+$ strains can secrete more pneumocins than BlpAB$^-$ strains during *com*-dependent *blp* activation in a time-dependent fashion (Fig. 2.12). The transient nature of ComAB expression during *com* activation limits pneumocin secretion in BlpAB$^-$ strains to short bursts following *com*-dependent *blp* activation. In contrast, *blp* activation – and therefore BlpAB expression – is not subject

to a rapid shut-off mechanism; in broth culture BlpR-regulated promoters remain highly active after initial activation throughout exponential phase (Fig. 2.4C, D). Thus, BlpAB$^+$ strains can sustain pneumocin secretion for longer periods of time and generate higher extracellular pneumocin concentrations due to secretion through BlpAB. Second, unlike BlpAB$^-$ strains, BlpAB$^+$ strains can activate *blp* independently of *com* (Fig. 2.8D, Fig. 2.10, (40)) with no penalty to pneumocin secretion capacity despite ComAB being unavailable to contribute to secretion (Fig. 2.12C).

Consistent with our in vitro findings, BlpAB$^+$ strains enjoy a competitive advantage over BlpAB$^-$ strains during nasopharyngeal colonization in mice. During co-colonization, BlpAB$^+$ strains were better able to outcompete a pneumocin-sensitive strain than BlpAB$^-$ strains (Fig. 2.13A). Moreover, a BlpAB$^+$ strain directly outcompeted its BlpAB$^-$ counterpart when both expressed the same pneumocins, but only when the two strains could not cross-activate each other's *blp* systems (Fig. 2.13B). These data indicate that compared to BlpAB$^-$ strains, BlpAB$^+$ strains can more effectively use pneumocins in competition during colonization due to more frequent activation of *blp* and/or secretion of greater amounts of pneumocins. Accordingly, we propose that BlpAB$^+$ strains leverage these abilities to act as aggressors, deploying pneumocins as weapons for generalized anti-bacterial competition. In contrast, BlpAB$^-$ strains primarily act as opportunists, using pneumocins in a limited capacity to augment fratricide and/or inhibit competitors while they are in a potentially vulnerable state during competence.

Linkage analysis indicates the predominant BlpAB-inactivating mutation moves from strain to strain via recombination and that BlpAB$^-$ strains occasionally revert to a BlpAB$^+$ genotype in the same manner (14). This suggests that there are costs and benefits to maintaining an intact BlpAB, and BlpAB$^+$ and BlpAB$^-$ strains represent dynamic populations that switch their BlpAB genotype in response to different selective pressures favoring an aggressor phenotype over opportunist or vice versa. Currently, genomics data do not indicate that invasion is such a selective pressure; in a collection of human isolates from South Africa (14), the distribution of BlpAB$^+$ strains did not differ between the invasive and colonizing groups (8/21 invasive vs. 11/30 colonizing, $p = 1$, Fisher's Exact Test). Another set of potential selective pressures is that associated with competence development. The increased propensity for competence activation in

BlpAB$^+$ strains (Fig. 2.8A, B, Fig. 2.10A) would provide them with greater access to DNA for horizontal gene transfer and an improved ability to cope with DNA-damaging stress. On the other hand, competence is a costly, energy-intensive process that interferes with normal cell metabolism and proliferation (84, 85). The stringent regulation of competence and its rapid shut-off once activated likely are strategies to mitigate its negative impacts. Therefore, the increased frequency of competence activation conferred by an intact BlpAB may be detrimental to fitness in certain cases. Finally, BlpAB$^+$ strains may incur a fitness cost from frequent *blp* activation. Consistent with this, we observed a dose-dependent, small but reproducible and statistically significant growth defect in THY broth-grown strains following BlpC treatment (Fig. 2.14). Validating these and other pressures in physiologic settings to further define how environmental and genetic contexts influence the adaptive value of maintaining an intact BlpAB presents an attractive target for future studies.

Many streptococci are naturally competent, but only members of the Mitis group (to which pneumococcus belongs) and the closely related Anginosus group (SAG) employ a ComCDE-type system to regulate competence (29). The ability of non-ComAB ABC transporters to secrete ComC/CSP and influence competence development is of direct relevance in other Mitis/Anginosus group species. For instance, *Sil* (*Streptococcus* invasion locus) is a bacteriocin-encoding locus (15) found in SAG that is structurally and functionally similar to pneumococcal *blp*. The regulation of *Sil* is effectively identical to that of *blp*, with SilE/D/CR/B/A taking the roles of BlpA/B/C/H/R, respectively. Importantly, SilED is the only ComAB/BlpAB-type transporter found in many SAG strains and therefore the only potential transporter for ComC/CSP. SAG SilE and SilCR/ComC share the same sequence motifs important for substrate recognition as those found in pneumococcal BlpA/ComA and BlpC/ComC, respectively (Fig. 2.15A).

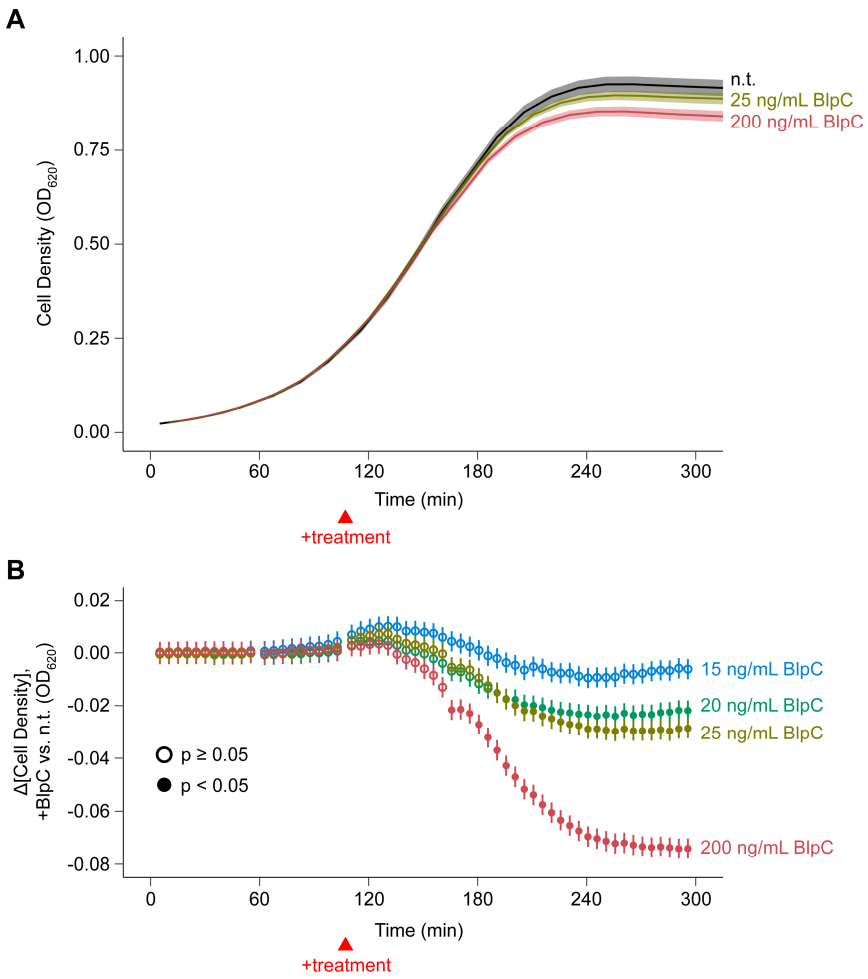

Figure 2.14. BlpC treatment induces a growth defect in broth culture. These data were obtained from the same set of experiments depicted in Figure 2.11B. **(A)** Growth curves of an R6-derived Δ*comC*Δ*blpC* mutant with and without BlpC treatment. Cells were grown in a 96-well plate in THY pH 7.1 and continuously monitored for cell density. Cells were either mock-treated (n.t.) or treated with the indicated concentrations of BlpC at t = 105 minutes. Data are graphed as the average (line) ± S.D. (shading) of 18 wells, pooled from 3 independent experiments. **(B)** BlpC-induced growth defect from panel A. Solid circles indicate a statistically significant defect ($p < 0.05$) at a particular timepoint. Data are graphed as the average ± S.E. of the difference between each BlpC-treated group and the mock-treated group. N = 18 wells per treatment, pooled from 3 independent experiments. Statistics: Linear mixed effects regression with treatment-control contrasts at each timepoint, Holm correction for multiple comparisons.

Therefore, it is likely that SilED can secrete CSP. Given this, competence in SAG may be chiefly regulated by a bacteriocin locus (*Sil*) through secretion of CSP through a non-ComAB transporter.

Looking beyond streptococci, members of the family of transporters to which ComAB and BlpAB belong, called PCATs, are widely distributed among gram-positive

and gram-negative bacteria, where they primarily function to export bacteriocins and in some cases also the signaling peptide that induces expression of said bacteriocins (4, 18, 19, 22, 35). With few exceptions, these peptides and transporters share the same conserved sequence motifs found to be important for substrate recognition in their pneumococcal counterparts (Fig. 2.15). This implies that most if not all ABC transporters of this family recognize their substrates not just in the same manner but also using the same sequence motifs. Therefore, the promiscuity of pneumococcal ComAB and BlpAB is likely a general feature of this transporter family. Consistent with this, others have shown that in *Lactococcus lactis*, the lactococcin Q transporter LaqD can secrete the related but distinct bacteriocin lactococcin G (22), and in *Enterococcus faecium*, the enterocin transporter EnkT recognizes the signal sequences of a number of bacteriocins from other species (38). This transporter promiscuity has wide-ranging implications for the regulation and biosynthesis of bacteriocins in many different

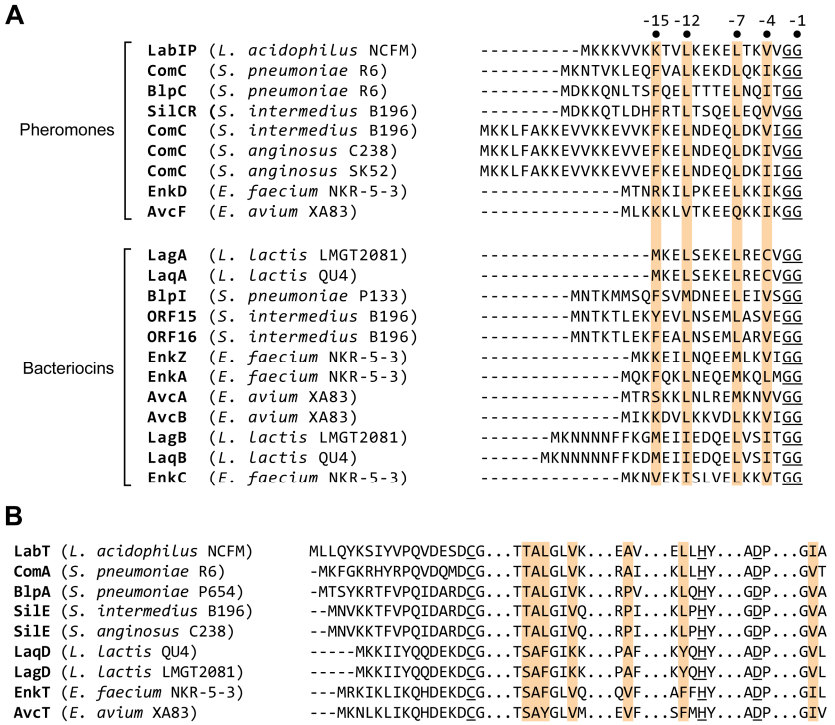

Figure 2.15. ComAB/BlpAB-type transporters and their double-glycine substrates share conserved sequence motifs. (A) Alignment of the N-terminal signal sequences of double-glycine peptides found in various species. The double-glycine motif is underlined. Residues predicted to interact with the transporter peptidase domain are highlighted in yellow. **(B)** Alignment of the peptidase domains of various ComA/BlpA-type transporters. The catalytic triad is underlined. Residues predicted to interact with substrates are highlighted in yellow.

bacterial species. Understanding the details of pheromone and bacteriocin secretion by these ABC transporters will provide key insights into the dynamics of inter-bacterial communication and competition.

2.5 Materials and Methods

2.5.1 Bacterial strains and growth conditions.

All strains are derived from the R6 strain P654 (183) (referred to as PSD100 in reference) or the D39 strain P2055 (Table 2.1). For experiments, pneumococcus was grown in either filter-sterilized THY (Todd Hewitt broth + 0.5% yeast extract) or CDM+ (105) at 37°C. Except where noted otherwise, pneumococcal cultures used for all experiments were inoculated from starter cultures grown in THY pH 7.4 to an OD_{620} of 0.275 and frozen at -80°C in 13% glycerol. Transformations were carried out as previously described (187). Unmarked chromosomal mutations were created via Janus or Sweet Janus exchange (188, 189). See Appendix A for details. All transformants were verified by PCR fragment length analysis and Sanger sequencing. Antibiotics were used at the following concentrations: chloramphenicol, 2 µg/mL; kanamycin, 500 µg/mL; spectinomycin, 200 µg/mL; streptomycin, 100 µg/mL.

Table 2.1. Strain list for Chapter II.

Strain	*Description*	*Antibiotic Resistances*	*Reference*
P654	R6, *rpsL*$_{167A>C}$, Δ*blpB-pncW*::[*blpB-pncW*]$_{6B}$, pE57 insertion in *blpA-blpQ* region (P$_{BIR}$-*lacZ*)	Str^R, Cam^R	(183)
P703	19A, *rpsL*$_{167A>C}$, BIR$_{P174}$, *blpH*$_{R6}$, *blpC*$_{R6}$, *blpA*$_{P174}$ (frameshifted), mouse passaged	Str^R	(14)
P824	19A, *rpsL*$_{167A>C}$, Δ*blp*::Janus, mouse passaged	Kan^R	Appendix A
P1163	19A, *rpsL*$_{167A>C}$, BIR$_{P164}$, *ant(9)-Ia* insertion 3' of *blpT*, mouse passaged	Str^R, Spc^R	Appendix A
P1613	P654 with pE57 insertion excised	Str^R	Appendix A
P1666	P1613 with pE135 insertion in *blpA-blpQ* region (P$_{BIR}$-*luc*)	Str^R, Cam^R	Appendix A
P2014	19A, *rpsL*$_{167A>C}$, BIR$_{P164}$, *blpA*$_{P174}$ (frameshifted), *ant(9)-Ia* insertion 3' of *blpT*, mouse passaged	Str^R, Spc^R	Appendix A
P2055	D39, *rpsL*$_{167A>C}$, Δ*blpT-blpA*::[*blpT-blpA*]$_{P32}$, ΔBIR::BIR$_{P164}$	Str^R	Appendix A
P2078	19A, *rpsL*$_{167A>C}$, BIR$_{P133}$, *blpH*$_{6A}$, *blpC*$_{6A}$, *ant(9)-Ia* insertion 3' of *blpT*, mouse passaged	Str^R, Spc^R	Appendix A
P2141	P654 with Δ*comC*::*comC-FLAG*	Str^R, Cam^R	Appendix A

Strain	*Description*	*Antibiotic Resistances*	*Reference*
P2166	P1613 with Δ*bgaA*::Janus-*luc*	Kan^R	Appendix A
P2187	P1613 with Δ*bgaA*::P_{comAB}-*luc*	Str^R	Appendix A
P2190	P654 with Δ*bgaA*::P_{comAB}-*Nluc*	Str^R, Cam^R	Appendix A
P2200	19A, $rpsL_{167A>C}$, BIR_{P133}, $blpH_{R6}$, $blpC_{R6}$, *ant(9)-Ia* insertion 3' of *blpT*, mouse passaged	Str^R, Spc^R	Appendix A
P2211	P1693 with Δ*comC*::*comC-FLAG*	Str^R, Cam^R	Appendix A
P2213	P1977 with Δ*comC*::*comC-FLAG*	Str^R, Cam^R	Appendix A
P2215	P2081 with Δ*comC*::*comC-FLAG*	Str^R, Cam^R	Appendix A
P2219	P2190 with P_{BIR}-*RFluc* in place of P_{BIR}-*lacZ*	Str^R, Cam^R	Appendix A
P2248	P2187 with Δ*comC*	Str^R	Appendix A
P2270	P2219 with Δ*comAB*	Str^R, Cam^R	Appendix A
P2272	P2219 with Δ*comC*	Str^R, Cam^R	Appendix A
P2287	P2219 with Δ*blpA* (*$blpA_{468\_469insAAGC}$*)	Str^R, Cam^R	Appendix A
P2309	P654 with Δ*blpQ-pncT*::*$blpI_{P133}$-FLAG*	Str^R, Cam^R	Appendix A
P2319	P2219 with Δ*blpC*	Str^R, Cam^R	Appendix A
P2321	P2270 with Δ*blpA* (*$blpA_{468\_469insAAGC}$*)	Str^R, Cam^R	Appendix A
P2329	P2309 with Δ*blpC*::*$blpC_{6A}$*	Str^R, Cam^R	Appendix A
P2341	P2329 with Δ*blpA* (*$blpA_{468\_469insAAGC}$*)	Str^R, Cam^R	Appendix A
P2343	P2329 with Δ*comAB*	Str^R, Cam^R	Appendix A
P2363	P2341 with Δ*comAB*	Str^R, Cam^R	Appendix A
P2384	P2363 with Δ*blpQ-pncT*::*$blpI_{P133}$-strepII*	Str^R, Cam^R	Appendix A
P2432	P1666 with Δ*comC*, Δ*blpC*	Str^R, Cam^R	Appendix A
P2475	P2329 with Δ*blpQ-pncT*::*$blpI_{P133}$-HiBiT*	Str^R, Cam^R	Appendix A
P2477	P2363 with Δ*blpQ-pncT*::*$blpI_{P133}$-HiBiT*	Str^R, Cam^R	Appendix A
P2483	P2341 with Δ*blpQ-pncT*::*$blpI_{P133}$-HiBiT*	Str^R, Cam^R	Appendix A
P2485	P2343 with Δ*blpQ-pncT*::*$blpI_{P133}$-HiBiT*	Str^R, Cam^R	Appendix A
P2516	P1666 with Δ*blpH*::*$blpH_{P1039}$*, Δ*comC*::*$comC_{TIGR4}$*, Δ*blpA* (*$blpA_{468\_469insAAGC}$*)	Str^R, Cam^R	Appendix A
P2518	P2516 with Δ*comAB*	Str^R, Cam^R	Appendix A
P2538	P2384 with Δ*comC*	Str^R, Cam^R	Appendix A
P2546	P2516 with Δ*comA*::*comA-TstrepII*	Str^R, Cam^R	Appendix A
P2563	P2475 with Δ*comC*::*$comC_{TIGR4}$*	Str^R, Cam^R	Appendix A
P2565	P2477 with Δ*comC*::*$comC_{TIGR4}$*	Str^R, Cam^R	Appendix A
P2567	P2483 with Δ*comC*::*$comC_{TIGR4}$*	Str^R, Cam^R	Appendix A
P2569	P2485 with Δ*comC*::*$comC_{TIGR4}$*	Str^R, Cam^R	Appendix A
P2604	P1666 with Δ*blpH*::*$blpH_{P1039}$*, Δ*comAB*, Δ*comC*, Δ*blpA* (*$blpA_{468\_469insAAGC}$*), Δ*blpC*	Str^R, Cam^R	Appendix A
P2610	P2055 with Δ*blpT-blpA*::[*blpT-blpA*]$_{P654}$, *$blpA_{1472A>G,1692C>A}$*, pE57-Δ*lacZ*::*RFluc* insertion (P_{BIR}-*RFluc* reporter)	Str^R, Cam^R	Appendix A
P2625	P2610 with Δ*bgaA*::P_{comAB}-*Nluc*	Str^R, Cam^R	Appendix A
P2665	P2625 with Δ*blpA*::*$blpA_{P654}$*	Str^R, Cam^R	Appendix A
P2666	P2625 with Δ*blpA* (*$blpA_{468\_469insAAGC}$*)	Str^R, Cam^R	Appendix A
P2668	P2625 with Δ*comAB*, Δ*blpA*::*$blpA_{P654}$*	Str^R, Cam^R	Appendix A
P2670	P2625 with Δ*comAB*, Δ*blpA* (*$blpA_{468\_469insAAGC}$*)	Str^R, Cam^R	Appendix A

2.5.2 DNA manipulation.

PCR for downstream Gibson assembly, transformation, or sequencing applications were performed using Phusion polymerase (NEB, E0553). All other PCR

reactions were performed using Taq polymerase (NEB, M0273). Primers (Table B.1) were designed using primer3 (190, 191) and synthesized by IDT. PCR products were purified using silica columns (Qiagen, 28106). Gibson assembly was performed using NEBuilder HiFi DNA Assembly master mix (NEB, E2621). All codon optimization was performed using OPTIMIZER (192).

2.5.3 com*/*blp *activation kinetics assays.*

Dual reporter strains were inoculated 1:150 into filter-sterilized THY + 25 mM HEPES, 5 µg/mL catalase, 165 µM D-luciferin (ThermoFisher, 88294), 1:10,000 Nano-Glo substrate (Promega, N113B) or CDM+ + 5 µg/mL catalase, 330 µM D-luciferin, 1:5,000 Nano-Glo substrate. Cultures were aliquoted into a white, clear-bottom 96-well plate (Porvair, 301012), 200 µL (if no treatment was to be added) or 198 µL (if treatment was to be added) per well. The plate was incubated at 37°C in a BioTek Synergy HTX plate reader. Every 5 minutes, the plate was shaken (linear shake setting, 587 cpm, 5 seconds) and absorbance (620 nm) and luminescence (bottom-read mode, 0.90-second integration time, gain 150) were read. Nluc signal was isolated using a 450/50 nm band-pass filter (BioTek, 7082208); RFluc signal was isolated using a 610 nm long-pass filter (BioTek, 7092209). For experiments in which pheromone treatments were added to the cells, 2 µL sterile media containing the appropriate pheromone (treatment) or 2 µL sterile media alone (mock treatment) were added to each well. Luminescence data were used to calculate *com*/*blp* activation levels (see the following section on activation kinetics calculations).

For analysis of spontaneous *com* and *blp* activation events, individual wells were assessed for activation events. An activation event was defined as when *com*/*blp* activation level exceeded a threshold value for at least two consecutive timepoints (see the following section on activation kinetics calculations). Wells without an activation event before cell density reached its maximum observed value were censored at the timepoint at which maximum cell density was reached. Once generated, activation event data were fit to the Kaplan-Meier estimator using the survfit() function in the R survival package (v2.41-3). The resulting “survival” curves were compared with the log-rank test using the survdiff() function in the R survival package (v2.41-3), and adjusted

for multiple comparisons with the Holm correction using the p.adjust() function in R 3.4.2.

2.5.4 com/blp *activation kinetics calculations.*

When NanoLuc (Nluc) and red *Luciola italica* firefly luciferase (RFluc) signals were simultaneously monitored, signal crosstalk from RFluc luminescence into the blue channel (450/50 BP) was not detectable. Signal crosstalk from Nluc luminescence into the red channel (610 LP) was minor and was corrected using the following formula:

$$L_i^{Rc} = L_i^R - k\left[L_i^B + \frac{L_{i+1}^B - L_i^B}{t_{i+1}^B - t_i^B}(t_i^R - t_i^B)\right]$$

L_i^{Rc}, corrected red channel luminescence at *i*th read; L_i^R, raw red channel luminescence at *i*th read; L_i^B, raw blue channel luminescence at *i*th read; L_{i+1}^B, raw blue channel luminescence at *i*+1th read; t_i^R, time of *i*th red channel luminescence read; t_i^B, time of *i*th blue channel luminescence read; t_{i+1}^B, time of *i*+1th blue channel luminescence read; k, cross-talk proportionality constant. The cross-talk proportionality constant, k, was empirically determined. A value of 0.0082 was used for the experiments in this study.

Luminescence values were converted to *com*/*blp* activation levels using the following formula:

$$A = \frac{\frac{dL}{dt} + k_d L}{N}$$

A, *com* or *blp* activation level; L, luminescence; t, time (min); N, cell density (OD_{620}), k_d, signal decay constant (min^{-1}). This formula models the levels of *com* and *blp* activation as the rate of transcription from ComE- and BlpR-regulated promoters, respectively. In the absence of post-transcriptional regulation of translation, the rate of transcription and rate of protein synthesis should be roughly linearly correlated. In turn, the rate of luciferase synthesis can be estimated from the first derivative of luminescence with respect to time ($\frac{dL}{dt}$, estimated from discrete data using the second-order central difference formula) corrected for the rate of signal decay, modeled as a first-order process. The signal decay constant, k_d, was empirically determined for Nluc (0.0462), RFluc (0.0347), and wild-type *Photinus pyralis* firefly luciferase (Luc) (0.0462).

Absorbance readings from the plate reader were corrected for pathlength to generate OD_{620} values (10-mm pathlength).

For detection of spontaneous *com* and *blp* activation events, *com* and *blp* activation levels were calculated as described above for individual wells in a 96-well plate. Analysis was restricted to timepoints at which OD_{620} was greater than 0.02. No activation events were observed at lower cell densities as judged by manual inspection of the data and the low values of both luminescence and OD_{620} at these early timepoints resulted in low signal-to-noise ratios that made automated analysis difficult. For both *com* and *blp*, an activation event was defined as the first timepoint at which the activation level remained above a threshold T for at least two consecutive readings. T was defined as a function of cell density as follows:

$$T = \frac{T^0}{\sqrt{N}}$$

T^0, threshold constant; N, cell density (OD_{620}). This function was chosen because it was the simplest form that empirically yielded a threshold curve with high sensitivity and specificity. The constant T^0 was empirically determined by manual inspection of activation level curves from pheromone-treated samples as positive controls and curves from non-treated pheromone deletion strains as negative controls.

2.5.5 Peptide processing assays.

R6 strains expressing FLAG-tagged peptides were inoculated 1:150 into THY + 25 mM HEPES pH 7.1, 5 µg/mL catalase and grown at 37°C statically. At OD_{620} 0.078, cells were induced with 100 ng/mL CSP1 and 100 ng/mL $BlpC_{R6}$. Thirty minutes later, a sample of cells, equivalent to OD_{620} 1.175 x 1 mL, was taken from each culture and pelleted by centrifugation, 5,000x*g*, 5 min, 4°C. The pellets were washed with 1 mL PBS and pelleted again using the same method. The washed pellets were then resuspended in 25 µL 1% SDS, 0.1% Triton X-100 and mixed with an equal volume of 4x Laemmli sample buffer (Bio-Rad, 1610747) supplemented with 10% (v/v) 2-mercaptoethanol. The resulting mixture was boiled at 95°C for 10 minutes, then stored at -20°C. The samples were later analyzed by peptide western blot (see the section on western blots). Blots were quantified using Image Studio v5.2. Percent peptide processing was calculated by dividing the signal from the processed peptide band by the sum of the

signals from the processed and unprocessed peptide bands. Experiments were repeated four times, and differences between groups were evaluated by applying ANOVA and Tukey's HSD to the percent peptide processing values using the aov() and TukeyHSD() functions in R 3.4.2.

2.5.6 BlpI-HiBiT secretion assays.

R6 strains expressing BlpI-HiBiT along with a background control strain expressing Strep-tag II-tagged BlpI were inoculated 1:150 into THY + 25 mM HEPES pH 7.1, 5 µg/mL catalase and grown at 37°C statically. These strains expressed a mismatched $BlpC_{6A}$-$BlpH_{R6}$ pair to prevent spontaneous *blp* activation. At OD_{620} 0.078, cells were induced with 200 ng/mL CSP1 and 200 ng/mL $BlpC_{R6}$. Sixty minutes later, a sample of culture supernatant of each strain was filter sterilized through a 0.22 µm centrifugal filter (Costar, 8160) and diluted ten-fold in sterile THY + 25 mM HEPES pH 7.1, 5 µg/mL catalase. Samples of mock-treated cultures were not diluted. At the same time, a sample of cells of each strain was taken and pelleted by centrifugation, 5,000x*g*, 5 min, 4°C. The pellet was washed once with 1x culture volume PBS and pelleted again by centrifugation, 5,000x*g*, 5 min, 4°C. The washed pellet was then resuspended in cell wall digest buffer [20 mM MES pH 6.5, 0.5 M sucrose, 20 mM $MgCl_2$, 50 U/mL mutanolysin (Sigma, M9901), 0.5 mg/mL lysozyme; adapted from (193)] and incubated at 37°C for 30 minutes. After incubation, the protoplast suspension was diluted 5-fold in 0.25% Triton X-100 and incubated at room temperature for 15 minutes.

For experiments simulating spontaneous activation, cells expressing mismatched CSP2-ComD1 and $BlpC_{6A}$-$BlpH_{R6}$ pairs were used. The background control strain was Δ*comC* in lieu of a mismatched CSP-ComD pair. For the single-timepoint assay, cells were induced starting at OD_{620} 0.078 as follows. "BlpC only": mock treatment followed by 25 ng/mL $BlpC_{R6}$ 45 minutes later; "simultaneous": mock treatment followed by 10 ng/mL CSP1 and 25 ng/mL $BlpC_{R6}$ 45 minutes later; "staggered": 10 ng/mL CSP1 followed by 25 ng/mL $BlpC_{R6}$ 45 minutes later. Supernatant and cell samples were collected 60 minutes after BlpC induction in all cases and processed as above. For the multiple-timepoint assay, cells were induced starting at OD_{620} 0.02 with either the "BlpC only" or "staggered" treatment as above. Supernatant samples were collected just before BlpC treatment (0 min) and every 30 minutes for 3 hours thereafter and were

clarified by centrifugation, 2750x*g*, 5 min, 4°C. All samples after the 0-minute timepoint were diluted 1:10 in sterile THY + 25 mM HEPES pH 7.1, 5 µg/mL catalase.

Standards were prepared by diluting synthetic L10-HiBiT (GGGGSGGGGSVSGWRLFKKIS; Genscript) in either 1:10 filter-sterilized background control strain supernatant (supernatant standards), undiluted filter-sterilized background control strain supernatant (supernatant standards for undiluted samples), or 1:5 background control strain lysate (cell lysate standards). For single-timepoint assays, the concentrations of L10-HiBiT used were 5-fold dilutions from 10 ng/mL to 640 fg/mL for the supernatant standards and 5-fold dilutions from 5 ng/mL to 320 fg/mL for the cell lysate standards. Standard curves for these were generated by fitting the log-transformed luminescence values for the standards against the log-transformed concentrations of each standard using the lm() function in R 3.4.2. For the multiple-timepoint assays, the concentrations of L10-HiBiT used were 10 ng/mL and 1 pg/mL for the diluted supernatant standards and 20 pg/mL and 200 fg/mL for the undiluted supernatant standards. Standard curves for these were generated by linear interpolation between the high and low standards. The limit of detection was defined as the concentration of L10-HiBiT in the lowest concentration standard used in the relevant standard curve. Wells that fell below the limit of detection were arbitrarily assigned a HiBiT concentration equal to half the limit of detection.

Supernatant and cell lysate samples of BlpI-HiBiT-expressing strains along with standards were aliquoted into a white 96-well plate (Costar, 3917), 50 µL per well, n=3 wells (single-timepoint assays), or 1 well (multiple-timepoint assays) per sample/standard. The samples and standards were then assayed for HiBiT tag by adding 50 µL HiBiT Extracellular Detection Reagent (Promega, N2421) to each well. Luminescence was then read using a BioTek Synergy HTX plate reader with the following settings: top read mode (1.00 mm height), 2-second integration time, gain 135. BlpI-HiBiT concentration in the samples were calculated using their luminescence values and the appropriate standard curve. Experiments were repeated three times and differences between groups were evaluated by applying ANOVA and Tukey's HSD to the log-transformed BlpI-HiBiT concentrations using the aov() and TukeyHSD() functions in R 3.4.2.

2.5.7 com/blp *activation characteristics assay.*

A Δ*comC* (*com* activation characteristics) or Δ*comC*Δ*blpC* (*blp* activation characteristics) R6 dual luciferase reporter strain was inoculated 1:150 into filter-sterilized THY + 25 mM HEPES pH 7.1, 5 µg/mL catalase, 165 µM D-luciferin (ThermoFisher, 88294), 1:10,000 Nano-Glo substrate (Promega, N113B). Cultures were aliquoted into a white, clear-bottom 96-well plate (Porvair, 301012), 198 µL (*com* activation characteristics) or 196 µL (*blp* activation characteristics) per well. The plate was incubated at 37°C in a BioTek Synergy HTX plate reader. The plate was monitored for absorbance and luminescence as described in the main text methods. When the average cell density reached approximately OD_{620} 0.078, 2 µL sterile THY alone or containing CSP1 was added to each well. For the *blp* activation characteristics assay, a second treatment consisting of 2 µL sterile THY alone or containing $BlpC_{R6}$ was added to each well 45 minutes after the first treatment. *com* and *blp* activation levels were calculated as described above. The area under the curve (AUC) of *com*/*blp* activation levels were calculated using trapezoidal sums.

2.5.8 ComA protein level kinetics assay.

R6 strains expressing Twin-Strep-tagged (194) ComA (ComA-TstrepII) or no ComAB (Δ*comAB*) and mismatched CSP2-ComD1 and $BlpC_{R6}$-$BlpH_{6A}$ pairs were inoculated 1:150 into THY + 25 mM HEPES pH 7.1, 5 µg/mL catalase, grown at 37°C statically, and induced with 10 ng/mL CSP1 at OD_{620} 0.078. Samples were taken just before induction (0 min) and 15, 30, 45, 75, and 105 minutes post-induction.

Membrane fractions were isolated using a protocol adapted from (193). Cell culture samples were pelleted by centrifugation, 5,000x*g*, 5 min, 4°C. The pellet was washed once with 1 mL PBS and pelleted again by centrifugation, 5,000x*g*, 5 min, 4°C. The washed pellet was then resuspended in 1 mL cell wall digest buffer [20 mM MES pH 6.5, 0.5 M sucrose, 20 mM $MgCl_2$, 100 U/mL mutanolysin (Sigma, M9901), 0.5 mg/mL lysozyme, 1:200 protease inhibitor cocktail III (Calbiochem, 539134)] and incubated at 37°C for 60 minutes. After incubation, the resulting protoplast suspension was pelleted by centrifugation, 3000x*g*, 10 min, room temperature. The supernatant (cell wall fraction) was removed, filter sterilized through a 0.22 µm centrifugal filter (Costar, 8160), and concentrated on an Amicon Ultra 30K filter (Millipore-Sigma,

UFC503024) to a factor of 13.3x. The pellet was frozen on dry ice for 5 minutes, resuspended in 1 mL protoplast lysis buffer [20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM DTT, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 10 µg/mL DNase I (Roche, 10104159001), 20 µg/mL RNase A (Millipore-Sigma, 70856), 1:200 protease inhibitor cocktail III], and incubated on ice for one hour with occasional mixing. The lysate was then pelleted by centrifugation, 16,300x*g*, 30 min, 4°C. The supernatant (soluble/cytoplasmic fraction) was removed, filter sterilized, and concentrated as above. The pellet was washed with 1 mL membrane wash buffer (20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM DTT, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 1:200 protease inhibitor cocktail III) and pelleted again by centrifugation, 16,300x*g*, 30 min, 4°C. The supernatant was discarded and the washed pellet (insoluble/membrane fraction) was solubilized in 75 µL 1% SDS + 0.1% Triton X-100 solution. 4x Laemmli sample buffer (Bio-Rad, 1610747) supplemented with 10% (v/v) 2-mercaptoethanol was added to all fractions to a final concentration of 1x. The fractions were then incubated at room temperature for 15 minutes and stored at -20°C.

Fractions were analyzed by western blot (see the section on western blots). Blots were quantified using Image Studio v5.2. Strep-tag II signal was normalized to total protein before being expressed as relative to the signal at the 15-minute timepoint.

2.5.9 Western blots.

Samples containing peptides for analysis were loaded onto 10-20% or 16.5% Tris-Tricine gels (Bio-Rad, 4563116 and 4563066) and run in Tris-Tricine SDS buffer (Bio-Rad, 1610744) at 100 V. All other types of samples were loaded onto 4-20% TGX gels (Bio-Rad, 4561096) and run in Tris-glycine SDS buffer (Bio-Rad, 1610732) at 200 V. Proteins and peptides were transferred to PVDF membranes (Millipore-Sigma, IPFL20200) using the wet transfer method in 25 mM Tris, 192 mM glycine, 20% methanol at 100 V, 1 hour, 4°C. For peptide westerns only, membranes were fixed after transfer using the method described in (195): membranes were dried at 37°C for 10 minutes, fixed in a 1:1 solution of 10% phosphate-buffered formalin (Fisher Scientific, 23-305510):PBS, shaking for 30 minutes at room temperature, and washed 3 x 5 min with PBS, shaking at room temperature. If necessary, total protein stain was performed after transfer using Revert Total Protein Stain Kit (Licor, 926-11010) according to manufacturer's instructions. Next, membranes were blocked with TBS (25 mM Tris-HCl

pH 7.5, 150 mM NaCl) + 5% skim milk, shaking for 1 hour at room temperature or overnight at 4°C. After blocking, membranes were washed 3 x 5 min with TBS-T (TBS + 0.1% Tween-20), shaking at room temperature. Then, membranes were incubated with primary antibody, shaking for 1 hour at room temperature or overnight at 4°C, then washed 3 x 5 min with TBS-T, shaking at room temperature. Membranes were then incubated with secondary antibody, shaking for 1 hour at room temperature, followed by 3 x 10 min washes with TBS-T, shaking at room temperature. Membranes were imaged using a Licor Odyssey scanner.

The following primary antibodies were used, listed along with their targets: FLAG, 1:1,000 mouse anti-FLAG M2 (Sigma, F3165) in TBS-T; Strep-tag II, 1:1,000 mouse anti-Strep-tag IgG2b (Qiagen, 34850) in TBS-T; pneumolysin, 1:2000 mouse anti-pneumolysin 1F11 (ThermoFisher, HYB-041-01-02) in TBS-T. The following secondary antibodies were used: 1:15,000 goat anti-mouse IgG IRDye 800CW (Licor, 925-32210) in TBS-T + 5% skim milk, 0.01% SDS; 1:15,000 goat anti-mouse IgG2b-specific IRDye 800CW (Licor, 926-32352) in TBS-T + 5% skim milk, 0.02% SDS.

2.5.10 Mouse nasopharyngeal colonization assays.

For the assays in which a BlpAB$^+$ or BlpAB$^-$ strain was competed against a whole-locus deletion Δ*blp* strain, P1163 and P2014 were used as the BlpAB$^+$ and BlpAB$^-$ “killer” strains, respectively, and P824 was used as the sensitive strain. P1163 and P2014 contain the BIR from P164 and express the two-peptide bacteriocin BlpIJ. For the assays in which pherotype matched and mismatched BlpAB$^+$/BlpAB$^-$ pairs were competed against each other, P2078 was used as the BlpCH$_{6A}$ BlpAB$^+$ strain, P2200 as the BlpCH$_{R6}$ BlpAB$^+$ strain, and P703 (14) (referred to as 19*Blp*$_{174}$ in reference) as the BlpCH$_{R6}$ BlpAB$^-$ strain. P2078 and P2200 contain the BIR from P133 and express the two-peptide bacteriocins BlpIJ and BlpMN as well as the bacteriocin BlpK. P703 contains the BIR (and frameshifted *blpA*) from P174. The genetic contents of BIR$_{P174}$ are the same as those of BIR$_{P133}$.

Mouse colonization was performed as previously described (14). Briefly, each strain was grown in THY to an OD$_{620}$ of 0.5. Cells were gently pelleted and resuspended in sterile PBS at 1/5 the original volume. For co-inoculations, PBS mixtures were combined in a 1:1 ratio; for single inoculations, an additional equal volume of sterile

PBS was added. 10 µL of dual or single-strain mixtures were inoculated into the nasopharynx of un-anaesthetized 5-7 week old female BALB/c mice (Taconic). Inocula were plated on selective media after inoculation to ensure that 1:1 ratios were maintained. No mouse developed evidence of sepsis during the experimental period. Mice were euthanized with CO_2 overdose on day 4 and nasopharyngeal colonization was sampled by cannulation of the trachea toward the nasopharynx and administration of 200 µL of sterile PBS. Samples were collected from the nares and used to obtain colony counts as described below.

Ten-fold serial dilutions, starting at 10^{-1}, of nasal washes were plated on trypticase soy agar (TSA) plates supplemented with 5 µg/mL catalase and the appropriate antibiotic(s), 100 µL per plate. Plates were incubated overnight at 37°C with 5% CO_2, after which colony counts were obtained. Counts from the lowest dilution yielding 30-500 colonies per plate were used to calculate colonization density (CFU/mL); if even the 10^{-1} dilution yielded less than 30 colonies, then the count from that dilution was used anyway. The limit of detection was set at 1000 CFU/mL, equal to 10 colonies on the 10^{-1} dilution plate. Samples that were below the limit of detection were arbitrarily assigned a colonization density equal to the limit.

For mice inoculated with a single strain, the following antibiotics were used for selection: spectinomycin for P1163, P2014, P2078, and P2200; kanamycin for P824; streptomycin for P703. For mice co-inoculated with P1163+P824 or P2014+P824, colonization density for individual strains were obtained by colony count as above following selection with spectinomycin for P1163 and P2014 and with kanamycin for P824. For mice co-inoculated with P2078+P703 or P2200+P703, colonization density for individual strains were obtained as follows. Total colonization density was obtained by colony count following selection with streptomycin. Colonization density of P2078 or P2200 was obtained by colony count following selection with spectinomycin. Colonization density of P703 was calculated by subtracting the P2078 or P2200 colonization density from the total colonization density.

Competitive index was calculated using the following formula:

$$CI = \frac{N_{1,out} N_{2,in}}{N_{1,in} N_{2,out}}$$

CI, competitive index; $N_{1,out}$, colonization density of strain 1; $N_{2,out}$, colonization density of strain 2; $N_{1,in}$, input CFU density of strain 1; $N_{2,in}$, input CFU density of strain 2. Colonization density was calculated as above. Input CFU density was determined by plating the initial inoculum and obtaining colony counts as above.

Mice were housed in a specific pathogen free facility and experiments were performed under an approved protocol in compliance with the University of Michigan Institutional Animal Care and Use Committee recommendations. Differences in colonization densities and competitive indices between groups were evaluated by the Mann-Whitney (2 groups) and Kruskal-Wallis (>2 groups) tests using the wilcox.test() and kruskal.test() functions in R 3.4.2.

2.5.11 BlpC-induced growth defect assay.

A Δ*comC*Δ*blpC* R6 dual luciferase reporter strain was monitored for cell density in response to BlpC treatment as described above in the section titled "*com*/*blp* activation characteristics assay". Only the wells that received mock treatment at the first treatment point (i.e. the wells that did not receive CSP) were analyzed. The data were fit to a linear mixed effects model using the lmer() function from the lme4 package (v1.1-17) in R 3.4.2. Time and BlpC treatment were modeled as discrete fixed effects and the date on which the experiment was performed as a random effect (random intercept model) to account for variation in the cell densities at which the treatment was added across different days. Differences between each BlpC treatment group and the mock-treated control group at each timepoint were assessed using the contrast() and lsmeans() functions in the emmeans package (v.1.1.3) in R 3.4.2, applying the Holm correction to the calculated p-values to adjust for multiple comparisons.

2.5.12 Nucleotide Accession Numbers.

The nucleotide sequences of P_{comAB}-Nluc, P_{BIR}-RFluc, and Sweet Janus+ are deposited in GenBank with accession numbers MH304212, MH304213, and MH304214, respectively.

2.6 Notes

The work presented in this chapter was published in the following article:

Wang CY, Patel N, Wholey WY, Dawid S. ABC transporter content diversity in *Streptococcus pneumoniae* impacts competence regulation and bacteriocin production. Proc Natl Acad Sci U S A. 2018 Jun 19;115(25):E5776-E5785.

Chapter III: Characterization of a Pneumococcal Rgg-regulated Double-Glycine Peptide Exporter Provides Insights into Molecular Determinants of Substrate Selectivity

3.1 Abstract

Peptidase-containing ABC transporters (PCATs) are a widely distributed family of transporters which secrete double-glycine (GG) peptides. In the opportunistic pathogen *Streptococcus pneumoniae* (pneumococcus), the PCATs ComAB and BlpAB have been shown to secrete quorum-sensing pheromones and bacteriocins related to the competence and pneumocin pathways. Here, we describe another pneumococcal PCAT, RtgAB, encoded by the *rtg* locus and found intact in 17% of strains. The Rgg/SHP-like quorum sensing system RtgR/S, which uses a peptide pheromone with a distinctive Trp-X-Trp motif, regulates expression of the *rtg* locus and provides a competitive fitness advantage in a mouse model of nasopharyngeal colonization. RtgAB secretes a set of co-regulated *rtg* GG peptides. ComAB and BlpAB, which share a substrate pool with each other, do not secrete the *rtg* GG peptides. Similarly, RtgAB does not efficiently secrete ComAB/BlpAB substrates. We examined the molecular determinants of substrate selectivity between ComAB, BlpAB, and RtgAB and found that the GG peptide signal sequences contain all the information necessary to direct secretion through specific transporters. Secretion through ComAB and BlpAB depends largely on the identity of four conserved hydrophobic signal sequence residues previously implicated in substrate recognition by PCATs. In contrast, a motif situated at the N-terminal end of the signal sequence, found only in *rtg* GG peptides, directs secretion through RtgAB. These findings illustrate the complexity in predicting substrate-PCAT pairings by demonstrating specificity that is not dictated solely by signal sequence residues previously implicated in substrate recognition.

3.2 Introduction

Export of polypeptides from their site of synthesis in the cytoplasm to the extracellular space is a fundamental physiological function for all cells. The secretome, the collection of all non-membrane associated proteins secreted from the cell, may comprise up to 20% of an organism's total proteome (1). Bacteria have evolved many different strategies for exporting proteins and peptides (2). One such strategy is the secretion of peptides using a family of ATP-binding cassette (ABC) transporters called peptidase-containing ABC transporters (PCATs).

PCATs are ABC transporters that contain characteristic N-terminal peptidase domains (PEP). PEP belongs to the family of C39 cysteine proteases and is responsible for the proteolytic processing of substrates during transport (4). In gram-positive bacteria, PCATs function either alone or with a single additional accessory protein (10). The most common function of PCATs is to assist in the biosynthesis of bacteriocins: antimicrobial peptides produced by bacteria to kill or otherwise inhibit the proliferation of other, usually closely related, bacteria (11). Some PCATs also promote cell-to-cell communication by secreting the peptide pheromones of gram-positive quorum-sensing systems (14, 15, 30, 86, 196). In short, PCATs are widely distributed peptide transporters which play key roles in shaping how bacteria interact with each other.

Oftentimes, expression of PCATs is under the control of quorum-sensing systems. These regulatory systems rely on cell-to-cell signaling to induce and coordinate the expression of their target genes under specific conditions. One such mode of PCAT regulation is the Rgg/SHP pathway (197). Rgg is a family of transcription regulators found in many gram-positive bacteria. In the genus Streptococcus, Rgg-family regulators are sometimes associated with short, hydrophobic peptides (SHPs) (198). SHPs are small peptides which are exported by the PptAB transporter (199, 200) and processed into mature pheromones. The Ami oligopeptide importer then internalizes the pheromones back into the cell, where they bind to and modulate the activity of Rgg-family regulators (198). Besides PCATs and bacteriocin production, Rgg/SHP systems have been found to regulate diverse processes such as carbohydrate utilization (201), tissue invasion (199, 202), capsule production (202, 203), and biofilm formation (32, 203). A related group of Rgg regulators, the ComRs, are

associated with SHP-like pheromones called ComS or XIP (sigX-inducing peptide) and control competence activation in some streptococcal species (204, 205).

In the gram-positive opportunistic pathogen *Streptococcus pneumoniae* (pneumococcus), the PCATs ComAB and BlpAB secrete quorum-sensing pheromones that control two important cellular pathways: genetic competence (the ability to take up and incorporate extracellular DNA into the genome) and production of the major family of pneumococcal bacteriocins (Blp bacteriocins, or pneumocins) (14, 24, 27, 86). ComAB and BlpAB secrete the same GG peptides, including the competence- and pneumocin-inducing pheromones and the pneumocins (39, 40, 206). Substrate sharing between ComAB and BlpAB affects competence and pneumocin regulation and influences when and with what effectiveness naturally occurring $BlpAB^{+}$ and $BlpAB^{-}$ strains can employ pneumocin-mediated killing (173, 206).

The functional implications of the shared ComAB/BlpAB substrate pool highlight the need to better understand how different PCATs select their substrates. PCAT substrates contain N-terminal signal sequences (also called leader peptides) which terminate in a Gly-Gly (sometimes also Gly-Ala or Gly-Ser) motif (4). For this reason, they are referred to as double-glycine (GG) peptides. During transport, PEP cleaves the peptide bond following the GG motif to remove the signal sequence from the C-terminal mature peptide fragment (cargo peptide). The signal sequences of GG peptides bind to PEP of PCATs through hydrophobic interactions involving three or four conserved residues in the signal sequences (7, 8, 36). These residues are located at positions −4, −7, −12, and −15 relative to the scissile bond. The GG motif allows the substrate to fit in the narrow entrance to the active site of PEP and is also required for binding and cleavage (5-7). Besides these conserved residues, the signal sequences of different GG peptides are fairly divergent. Mutagenesis studies of several different PCAT-substrate pairs have largely failed to identify any contribution of these non-conserved residues to substrate-PEP binding (7, 8, 36).

While substantial progress has been made in uncovering the mechanisms that allow PCATs to recognize GG peptides, comparatively little is known about how or if PCATs discriminate between different GG peptides. In addition to ComAB and BlpAB from pneumococcus, multiple PCATs have been shown to process and/or secrete

multiple peptides with distinct signal sequences, sometimes even those from different strains or species (7, 18, 21, 22, 38). These data suggest that in general, PCATs are not particularly selective when it comes to choosing substrates.

In this work, we describe a previously uncharacterized locus in pneumococcus, *rtg*, which encodes the PCAT RtgAB and several GG peptides. This locus is regulated by the Rgg/SHP-like system RtgR/S, which provides a competitive fitness advantage during nasopharyngeal colonization. We demonstrate that RtgAB secretes the *rtg* GG peptides but not ComAB/BlpAB substrates, nor can ComAB or BlpAB efficiently secrete the *rtg* GG peptides. Finally, we investigate the signal sequence determinants that selectively direct peptides toward either RtgAB or ComAB/BlpAB and show that a unique N-terminal motif is required for secretion by RtgAB. These findings shed light on how PCATs can use signal sequence motifs beyond the previously described conserved hydrophobic residues to distinguish different GG peptides.

3.3 Results

3.3.1 Identification of an uncharacterized pneumococcal PCAT-encoding locus.

As part of an effort to catalog the PCAT repertoire of *S. pneumoniae*, we searched pneumococcal genomes for putative PCAT genes that had not been previously described. One of the hits was *CGSSp9BS68_07257* (henceforth *07257*), a gene found in the clinical isolate Sp9-BS68 (207) (Fig. 3.1A). Upstream of *07257* is a gene oriented in the opposite direction predicted to encode an Rgg-family transcription regulator (198). We hypothesized that this regulator controls expression of 07257 and named the locus *rtg* (Rgg-regulated transporter of double-glycine peptides), the transporter gene *rtgA* and the regulator gene *rtgR*. *rtgR* marks one end of the locus and is separated from a partially disrupted arginine biosynthesis cluster (208) by two transcription terminators. Downstream of *rtgA* are several genes arranged in a single operon. These include *rtgB*, which encodes a putative ComB/BlpB-like transport accessory protein, and the GG peptide genes *rtgG*, *rtgT*, *rtgW1* and *rtgW2*. A transcription terminator separates the last gene, *rtgD2*, from a disrupted putative endoRNase gene and *pspA*. A different version of the *rtg* locus is also found in the laboratory strain D39 with a disrupted *rtgA* (Fig. 3.1A).

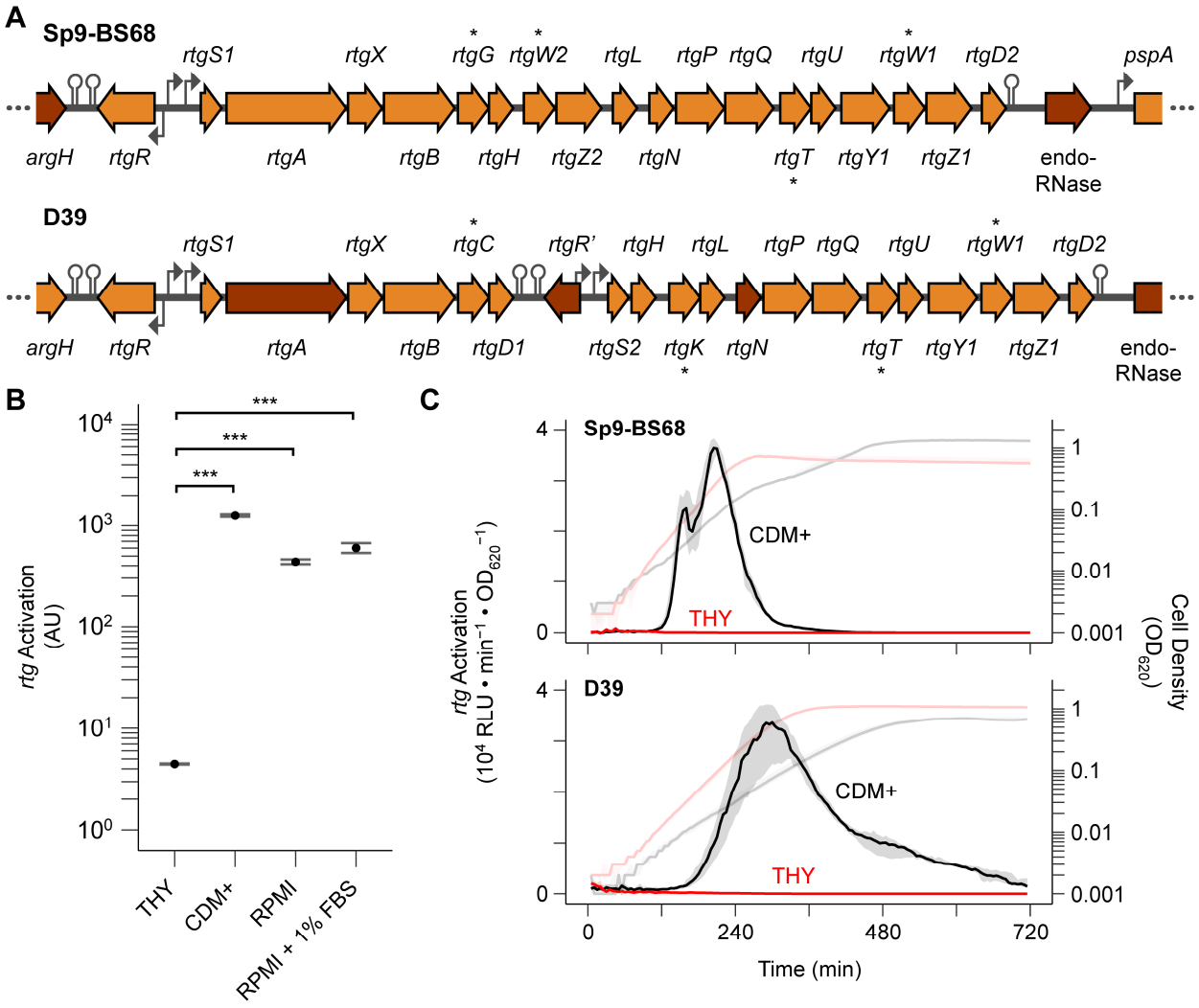

Figure 3.1. *rtg* is an actively regulated PCAT-encoding locus in pneumococcus. (A) Organization of *rtg* in strains Sp9-BS68 and D39. Block arrows represent genes. Dark shading indicates a pseudogene. Bent arrows represent promoters. Hairpins represent transcription terminators. GG peptide genes are marked with asterisks. **(B)** Activation of *rtg* in various growth media. A Sp9-BS68 P_{rtgA}-*luc* reporter strain was grown in the indicated media, and luciferase activity was sampled when cells reached OD_{620} 0.2. The plotted values were normalized against cell density and the luciferase activity from a strain harboring constitutively expressed luciferase grown in the same medium. Plots show mean ± S.E. of 3 independent experiments. Statistics: *** $p < 0.001$; ANOVA with Tukey's HSD. **(C)** Timing of *rtg* activation in THY and CDM+. Sp9-BS68 (top) and D39 (bottom) P_{rtgS1}-*luc* reporter strains were grown in THY (red) and CDM+ (black) and monitored for *rtg* activation (dark, left y axis) and cell density (light, right y axis). Plots show median (line) and 25% to 75% quantiles (shading) of 12 wells pooled from 3 independent experiments.

3.3.2 The Rgg/SHP-like pheromone pair RtgR/RtgS regulates rtg*.*

We found that *rtg* expression is inhibited in mid-exponential phase during growth in the peptide-rich medium THY (Todd-Hewitt broth + 0.5% yeast extract) but highly

upregulated in the peptide-poor media CDM+ (105) and RPMI (Fig. 3.1B). The start of *rtg* activation in both Sp9-BS8 and D39 during growth in CDM+ occurs in early exponential phase at cell densities as low as OD_{620} 0.01 (Fig. 3.1C). In contrast, *rtg* stays inactive in THY throughout the exponential and stationary phases. We concluded from these data that *rtg* is actively regulated, most likely by RtgR. Since Rgg regulators are often associated with peptide pheromones, we searched for and found an ORF in Sp9-BS68 between *rtgR* and *rtgA* predicted to encode a SHP-like pheromone. D39 has two copies of the candidate pheromone: one located between *rtgR* and *rtgA* and the other downstream of *rtgB*. We named the only copy of the ORF in Sp9-BS68 and the first copy in D39 (between *rtgR* and *rtgA*) *rtgS1* and the second copy in D39 *rtgS2* (Fig. 3.1A).

Having identified a putative Rgg/SHP-like regulatory system, we sought to define the contributions of RtgR and RtgS to *rtg* regulation through deletional analysis. We monitored *rtg* activation in Sp9-BS68 Δ*rtgS1*, Δ*rtgR*, and Δ*rtgS1*Δ*rtgR* strains during growth in CDM+ and THY (Fig. 3.2A). None of the mutants showed signs of *rtg* activation in either medium, indicating that both RtgR and RtgS promote and are required for *rtg* activation. In D39, *rtgS1* encodes a peptide with a single amino acid change (S14L) compared to RtgS from Sp9-BS68, and *rtgS2* encodes a peptide with a different single amino acid change (P27S) (Fig. 3.2B). We found that the D39 Δ*rtgS1* and Δ*rtgS1*Δ*rtgS2* mutants failed to activate *rtg* in CDM+ while the Δ*rtgS2* mutant was indistinguishable from the wild-type strain (Fig. 3.2C). This indicated that the P27S substitution in the *rtgS2* product prevents it from activating *rtg*, while the S14L substitution in the *rtgS1* product does not appreciably affect signaling. Therefore, we classified the *rtgS1* product in both Sp9-BS68 and D39 as type A pheromone ($RtgS_A$) and the *rtgS2* product in D39 as type B ($RtgS_B$).

To confirm that RtgS is the specific pheromone inducer of *rtg*, we performed dose-response assays using synthetic peptides corresponding to the C-terminal eight, ten, and twelve residues of $RtgS_A$ ($RtgS_A$-C8, $RtgS_A$-C10, $RtgS_A$-C12, respectively). All three synthetic peptides induced expression from the Sp9-BS68 *rtgS1* promoter in both CDM+ and THY, though the curves for the latter were shifted to the right in a manner consistent with pure competitive inhibition (Fig. 3.2D, Table 1). We also confirmed that

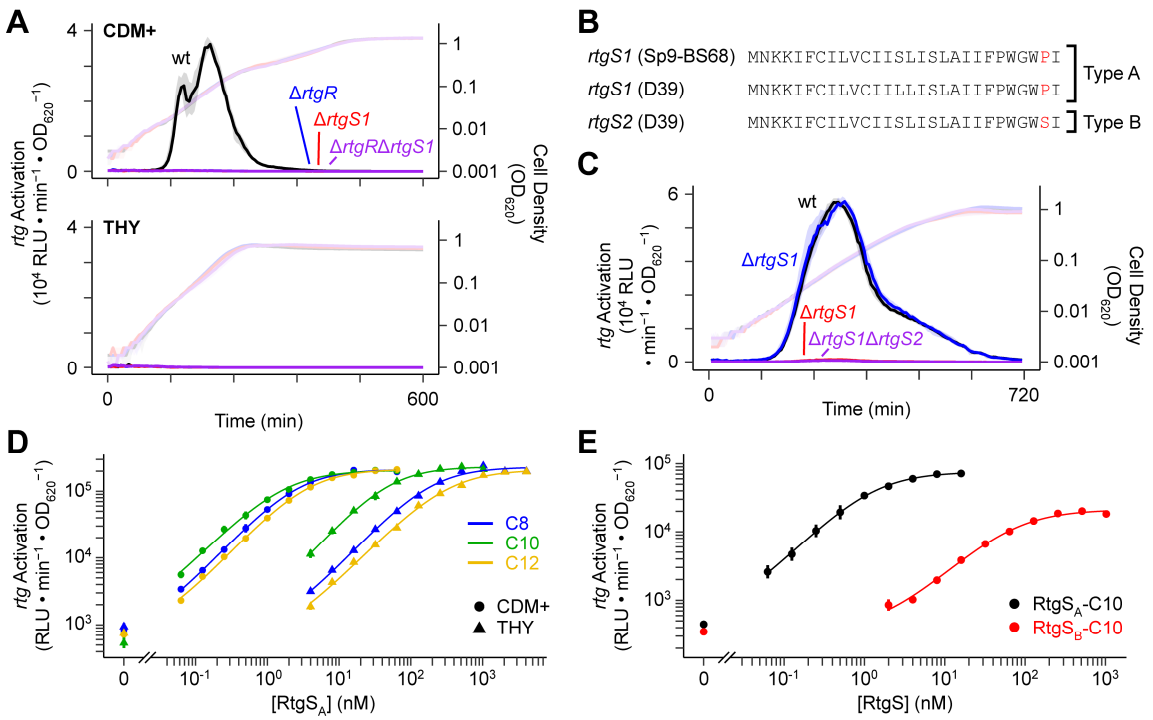

Figure 3.2. The Rgg/SHP-like RtgR/RtgS system regulates *rtg*. (A) *rtgR* and *rtgS1* are required for *rtg* activation in Sp9-BS68. Sp9-BS68 P_{rtgS1}-*luc* reporters were grown in CDM+ or THY and monitored for *rtg* activation (dark, left y axis) and cell density (light, right y axis). Plots show median (line) and 25% to 75% quantiles (shading) of 12 wells pooled from 3 independent experiments. **(B)** Translated *rtgS* gene products from Sp9-BS68 and D39. The type-defining residue is highlighted in red. **(C)** *rtgS1* but not *rtgS2* is required for *rtg* activation in D39. D39 P_{rtgS1}-*luc* reporters were grown in CDM+ and monitored for *rtg* activation (dark, left y axis) and cell density (light, right y axis). Plots show median (line) and 25% to 75% quantiles (shading) of 30 wells pooled from 3 independent experiments. **(D)** C-terminal fragments of $RtgS_A$ induce *rtg*. A Sp9-BS68 Δ*rtgS1* P_{rtgS1}-*luc* reporter was grown in CDM+ or THY to OD_{620} 0.02 and treated with synthetic $RtgS_A$ fragments. Response was defined as the maximum observed P_{rtgS1} activity within 60 min of treatment. Plotted data points represent mean ± S.E. of 3 independent experiments. **(E)** $RtgS_B$ is a partial agonist at the *rtg* locus. A D39 Δ*rtgS1*Δ*rtgS2* P_{rtgS1}-*luc* reporter was grown in CDM+ to OD_{620} 0.02 and treated with synthetic RtgS fragments. Response was defined as the maximum observed P_{rtgS1} activity within 60 min of treatment. Plotted data points represent mean ± S.E. of 3 independent experiments.

rtg induction by synthetic RtgS requires RtgR (Fig. 3.3). In D39, $RtgS_A$-C10 induces *rtg* similarly to Sp9-BS68 whereas $RtgS_B$-C10 acts as a partial agonist with a 55-fold larger EC50 value than $RtgS_A$-C10 (Fig. 3.2E, Table 3.1). Therefore, the Pro to Ser substitution in $RtgS_B$ interferes with signaling at a step following pheromone secretion. Although partial agonists can act as competitive antagonists of full agonists, we did not observe an inhibitory phenotype associated with $RtgS_B$ during natural *rtg* activation (Fig. 3.2C) and competitive dose-response assays showed that $RtgS_B$-C10 only antagonizes $RtgS_A$-C10 at likely supraphysiological concentrations (≥ 256 nM) (Fig. 3.4).

Table 3.1. RtgS dose-response parameters.

Strain	*Growth Medium*	*Peptide*	*EC50 (nM)*[a]	*Maximal response* (10^3 *RLU* · min^{-1} · OD_{620}^{-1})[a]
Sp9-BS68[b]	CDM+	$RtgS_A$-C8	2.53 ± 0.25	212.2 ± 9.4
		$RtgS_A$-C10	1.63 ± 0.19	205.7 ± 9.1
		$RtgS_A$-C12	3.51 ± 0.28	213.9 ± 8.0
	THY	$RtgS_A$-C8	184 ± 15	229.7 ± 8.6
		$RtgS_A$-C10	47.4 ± 6.9	232.4 ± 14.3
		$RtgS_A$-C12	315 ± 42	210.1 ± 13.1
D39[c]	CDM+	$RtgS_A$-C10	1.25 ± 0.30	75.1 ± 8.1
		$RtgS_B$-C10	68.8 ± 11.5	21.2 ± 1.5

[a] Mean ± S.E.; derived from fitting data to Hill model.
[b] Δ*rtgS1*
[c] Δ*rtgS1*Δ*rtgS2*

Studies of pneumococcal Rgg/SHP systems have yet to experimentally identify the transporters responsible for pheromone export and import (32, 202, 203). In Rgg/SHP systems found in non-pneumococcal streptococci, the PptAB and Ami transporters carry out pheromone export and import, respectively (198-200). In order to determine if these two transporters perform the same functions with respect to RtgS, we assayed the ability of the Sp9-BS68 Δ*pptAB* and Δ*amiCD* mutants to activate *rtg*. Neither mutant spontaneously activates *rtg* in CDM+ (Fig. 3.5A). This *rtg* activation defect could be fully rescued with exogenous RtgS treatment in the Δ*pptAB* mutant but not the Δ*amiCD* mutant (Fig. 3.5B). These data show that the *pptAB* deletion blocks a step in the RtgS life cycle prior to pheromone import and that the *amiCD* deletion blocks a step after pheromone export and maturation. This is consistent with a model in which PptAB exports RtgS and Ami imports RtgS.

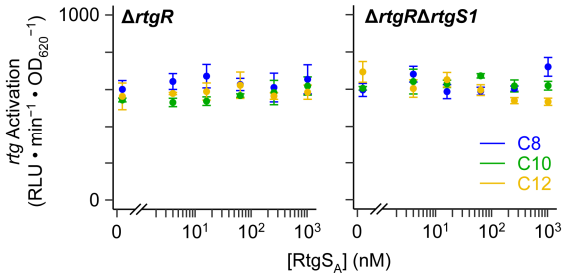

Figure 3.3. RtgS-induced *rtg* activation requires RtgR. Sp9-BS68 Δ*rtgR* and Δ*rtgR*Δ*rtgS1* P_{rtgS1}-*luc* reporters were grown in CDM+ to OD_{620} 0.02 and treated with synthetic $RtgS_A$ fragments. Response was defined as the maximum observed P_{rtgS1} activity within 60 min of treatment. Plotted data points represent mean ± S.E. of 3 independent experiments.

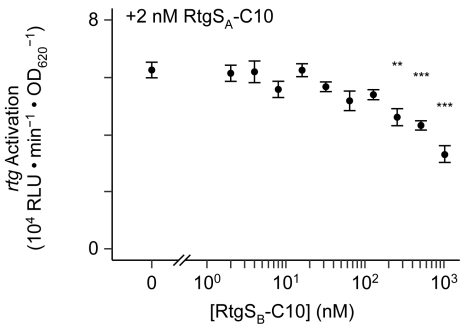

Figure 3.4. High concentrations of RtgS$_B$-C10 antagonize RtgS$_A$-C10. D39 Δ*rtgS1*Δ*rtgS2* P$_{rtgS1}$-*luc* reporters were grown in CDM+ to OD_{620} 0.02 and treated simultaneously with 2 nM RtgS$_A$-C10 and various concentrations of RtgS$_B$-C10. Response was defined as the maximum observed P$_{rtgS1}$ activity within 120 min of treatment. Plotted data points represent mean ± S.E. of 3 independent experiments. Statistics: comparisons vs. 0 nM RtgS$_B$-C10; ** $p < 0.01$, *** $p < 0.001$; ANOVA with Dunnett's correction for multiple comparisons.

Finally, we characterized the promoter regions of *rtgR* and *rtgS1* in Sp9-BS68. We found one putative promoter for *rtgR* consisting of an extended −10 element (Fig. 3.6A). We also found two putative promoters for *rtgS1* (P1 and P2), each containing a −10 element downstream of two overlapping imperfect inverted repeats (IR1 and IR2) (Fig. 3.6A). A luciferase reporter driven only by P2 showed roughly 30% less luciferase activity compared to one driven by both P1 and P2 (Fig. 3.6B), indicating that both promoters contribute to the expression of the *rtgS1* operon. Because Rgg-family and

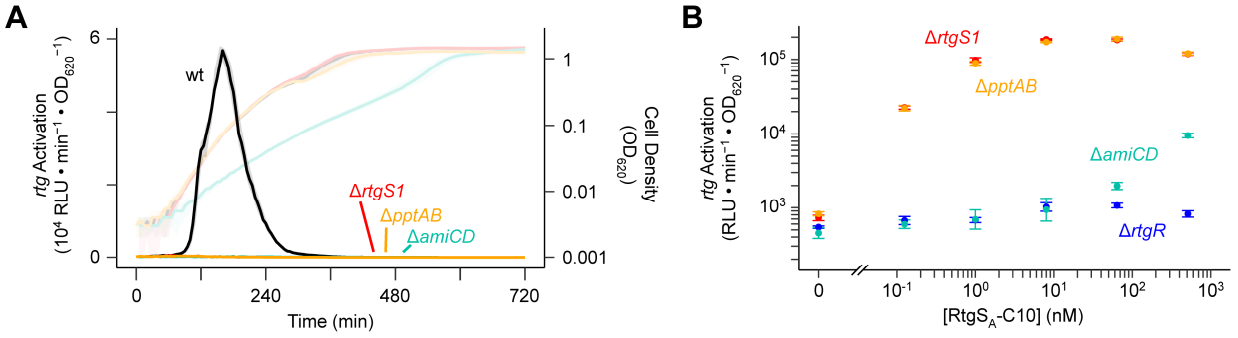

Figure 3.5. Ami and PptAB are required for RtgR/S signaling. (A) Both Ami and PptAB are required for natural *rtg* autoinduction. Sp9-BS68 P$_{rtgS1}$-*luc* reporters were grown in CDM+ and monitored for *rtg* activation (dark, left y axis) and cell density (light, right y axis). Plots show median (line) and 25% to 75% quantiles (shading) of 30 wells per strain pooled from 3 independent experiments. **(B)** Exogenous RtgS treatment rescues *rtg* activation defect in the Δ*pptAB* mutant but not the Δ*amiCD* mutant. Sp9-BS68 P$_{rtgS1}$-*luc* reporters were grown in CDM+ to OD_{620} 0.02 and treated with RtgS$_A$-C10. Response was defined as the maximum observed P$_{rtgS1}$ activity within 60 min of treatment. Plotted data points represent mean ± S.E. of 3 independent experiments.

related regulators can bind to inverted repeats (205, 209-211), we hypothesized that IR1 and/or IR2 are RtgR binding sites. To test this hypothesis, we determined the effect of various nucleotide substitutions in P2 (Fig. 3.6A) on RtgS-induced expression from this promoter. Substitutions in IR1 render the promoter insensitive to RtgS treatment, while substitutions immediately upstream of IR1 and in IR2 only modestly affect expression in either direction (Fig. 3.6C). These data are consistent with a model in

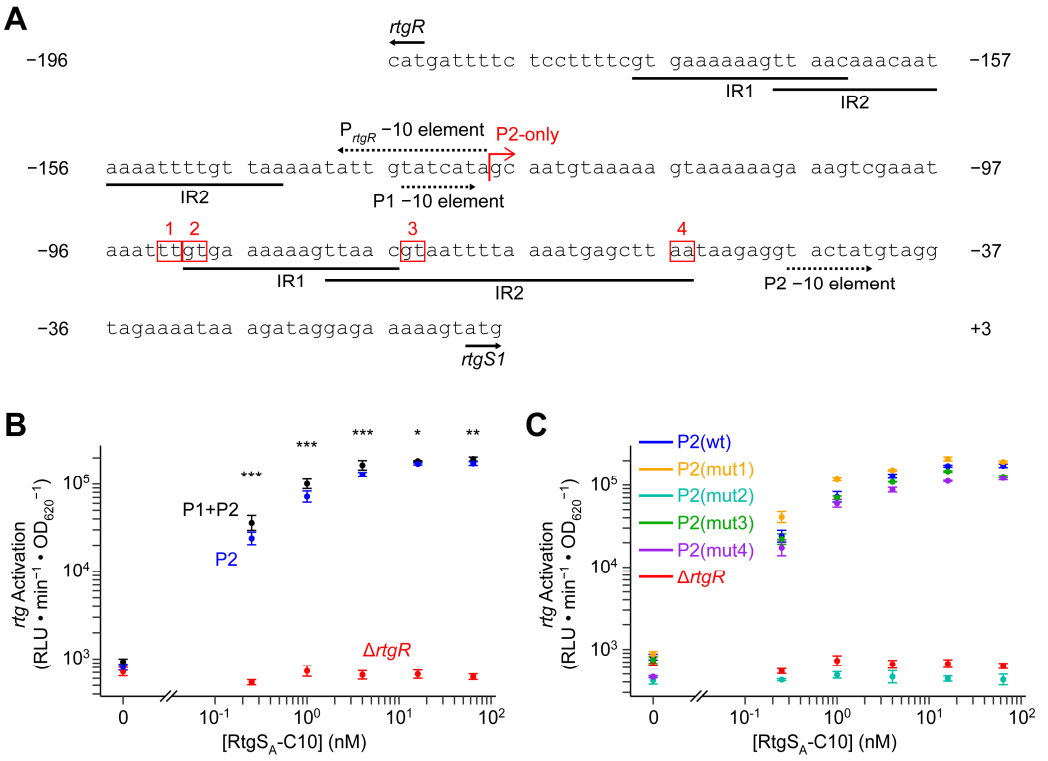

Figure 3.6. Two promoters, P1 and P2, both with putative RtgR binding sites, contribute to *rtgS1* expression. (A) Promoter region of *rtgR* and *rtgS1* in Sp9-BS68. Inverted repeats (IR) are underlined. −10 elements are marked with dotted arrows. Start sites of *rtgR* and *rtgS1* are marked with solid arrows. Red bent arrow indicates span of P2 promoter construct used in panel B. Numbered, red boxes indicate mutated nucleotides used in panel C. **(B)** P1 and P2 each partially contribute to *rtgS1* expression. Sp9-BS68 Δ*rtgS1* P_{rtgS1}-*luc* reporters were grown in CDM+ to OD_{620} 0.02 and treated with synthetic $RtgS_A$-C10. Response was defined as the maximum observed P_{rtgS1} activity within 60 min of treatment. Plotted data points represent mean ± S.E. of 3 independent experiments. Statistics: comparisons between P1+P2 and P2 reporters at each dose; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; ANOVA followed by Holm-corrected post-tests. **(C)** IR1 is required for RtgS-induced *rtg* activation. Sp9-BS68 Δ*rtgS1* P_{rtgS1}-*luc* reporters were grown in CDM+ to OD_{620} 0.02 and treated with synthetic $RtgS_A$-C10. Except for the Δ*rtgR* strain, *luc* in each reporter was fused to P2 only. Response was defined as the maximum observed P_{rtgS1} activity within 60 min of treatment. Plotted data points represent mean ± S.E. of 3 independent experiments.

which RtgR binds IR1 to upregulate transcription from P1 and P2 in response to RtgS signaling.

3.3.3 RtgAB secretes rtg*-encoded GG peptides.*

Sp9-BS68 and D39 both encode four putative GG peptides at the *rtg* locus (Fig. 3.1A). To determine whether RtgAB secretes the *rtg* GG peptides, we employed a HiBiT tag-based peptide secretion assay (206). We constructed autoinducing-deficient Δ*rtgS1*Δ*rtgS2* strains in the R6 background expressing HiBiT-tagged RtgC (from D39) or RtgG (from Sp9-BS68) from downstream of *rtgB*. These strains were either ComAB^{-}/BlpAB^{-}/ RtgAB^{-} or ComAB^{-}/BlpAB^{-}/ RtgAB^{+}. Upon RtgS-C10 treatment in CDM+, levels of extracellular RtgC-HiBiT and RtgG-HiBiT in the RtgAB^{+} cultures

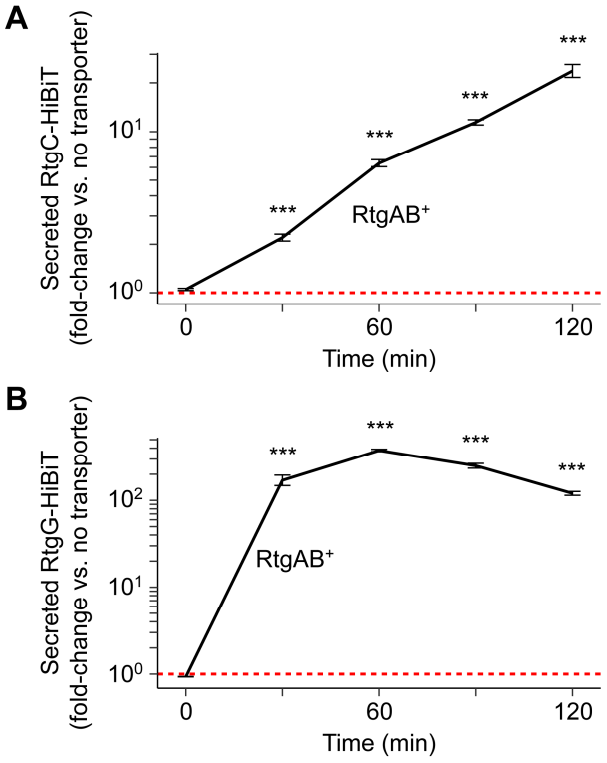

Figure 3.7. RtgAB secretes RtgC and RtgG. (A, B) R6 ComAB^{-}/BlpAB^{-}/RtgAB^{-} and ComAB^{-}/BlpAB^{-}/RtgAB^{+} strains expressing RtgC-HiBiT (A) or RtgG-HiBiT (B) were grown in CDM+ to OD_{620} 0.05 and treated with 200 ng/mL CSP, 200 ng/mL BlpC, and 20 nM $RtgS_A$-C10. Samples were taken every 30 min and extracellular HiBiT signal was quantified. Data are presented as fold-change values relative to the ComAB^{-}/BlpAB^{-}/RtgAB^{-} control. Red, dashed line represents fold-change = 1 (no difference vs. the control). Plots show mean ± S.E. of 3 independent experiments. Statistics: comparisons vs. ComAB^{-}/BlpAB^{-}/RtgAB^{-} at each timepoint; *** $p < 0.001$; ANOVA with Tukey's HSD.

Table 3.2. Genes of the *rtg* locus.

Gene		*Product*	*Frequency*[a,b]
rtgA	(intact)	Peptidase-containing ABC transporter	5.0%
	(start codon mutation only)	Peptidase-containing ABC transporter	14.5%
	(disrupted)	Peptidase-containing ABC transporter, truncated	80.5%
rtgB	(intact)	Putative transport accessory protein	18.6%
	(disrupted)	Putative transport accessory protein, truncated	2.5%
rtgC		GG peptide	15.1%
rtgD		Hypothetical protein	99.7%
rtgE		GG peptide	7.2%
rtgG		GG peptide	4.1%
rtgH		Hypothetical protein	40.6%
rtgK		GG peptide	41.2%
rtgL		Hypothetical protein	83.3%
rtgM		GG peptide	59.7%
rtgN		Hypothetical protein	82.4%
rtgP		Hypothetical protein	83.3%
rtgQ		Hypothetical protein	83.3%
rtgR		Rgg-family transcription regulator	78.6%
rtgR'		Rgg-family transcription regulator, truncated	36.5%
rtgS	(type A)	SHP-like pheromone, type A	42.5%
	(type B)	SHP-like pheromone, type B	67.6%
	(type C)	SHP-like pheromone, type C	0.6%
rtgT		GG peptide	83.0%
rtgU		Hypothetical protein	83.0%
rtgV		GG peptide; probable fusion of RtgW and RtgZ	35.5%
rtgW		GG peptide	83.0%
rtgX		GyrI-like hypothetical protein	21.1%
rtgY		Hypothetical protein	88.7%
rtgZ		Hypothetical protein	99.7%

[a] In 318 genomes from the Massachusetts collection with fully sequenced, gap-less *rtg* loci.
[b] May include pseudogenes unless stated otherwise.

increase to 26- and 376-fold, respectively, compared to their levels in the RtgAB$^-$ cultures (Fig. 3.7). From these data, we conclude that RtgAB secretes the *rtg* GG peptides.

3.3.4 The rtg *locus exhibits extensive variation across different pneumococcal strains.*

In order to catalog the diversity found at *rtg*, we conducted a survey of the locus in a collection of pneumococcal clinical isolates from Massachusetts, USA (212). After removing genomes in which *rtg* spans multiple contigs or conserved genes flanking *rtg* could not be found, we were left with 318 out of 616 strains, all of which encoded at least one *rtg* gene. We analyzed the *rtg* loci from these 318 strains and clustered them into 23 groups based on overall architecture (Fig. 3.8). Across all 318 loci, we found 24 unique *rtg* genes (Table 3.2). Searching for these genes in the full collection revealed

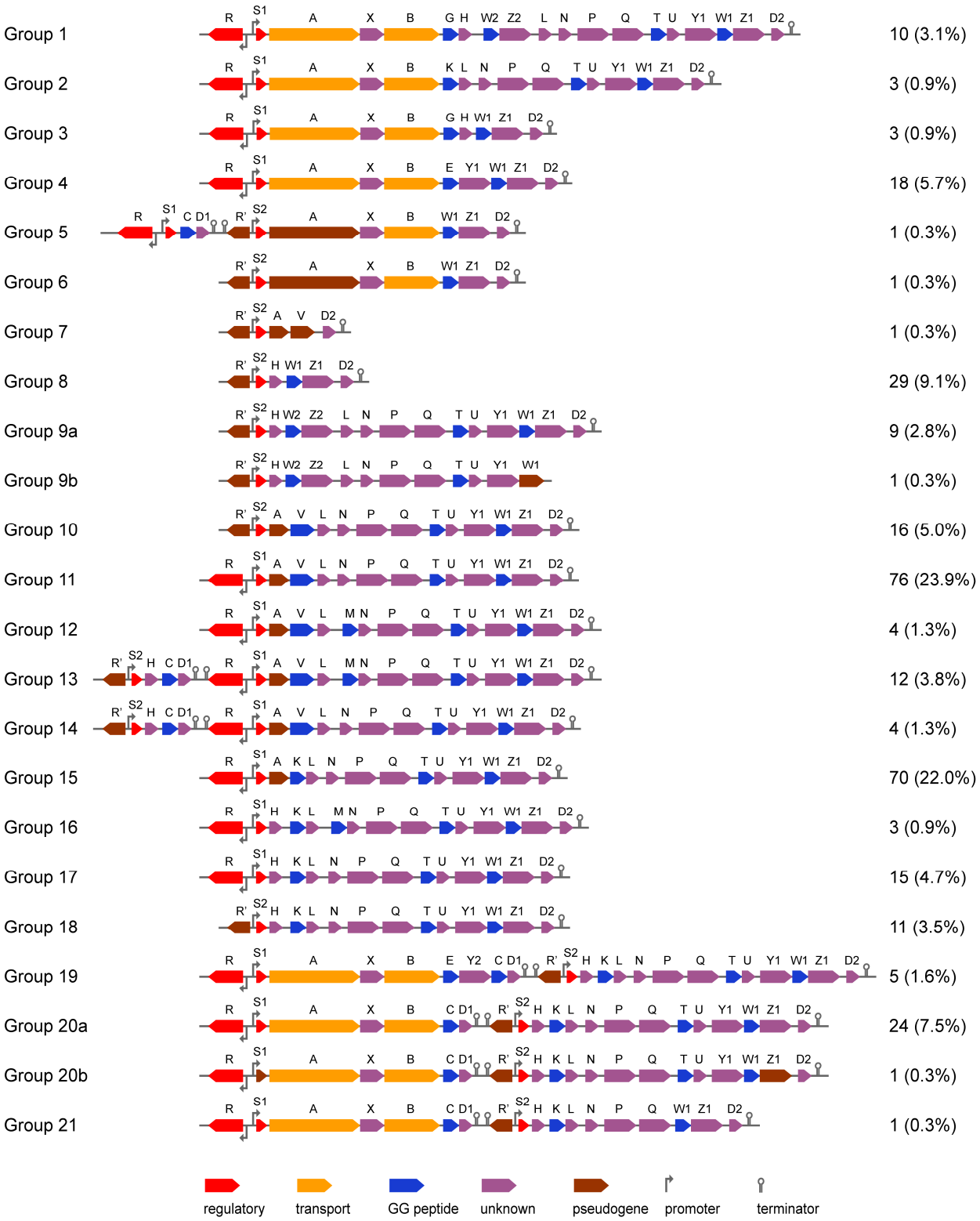

Figure 3.8. Diversity of fully sequenced *rtg* loci in the Massachusetts pneumococcal isolate collection. *rtg* loci from a filtered set of 318 strains from the Massachusetts isolate collection with gapless sequence coverage of the *rtg* genomic locale were clustered into 23 groups based on gene content and synteny. The organization of *rtg* in each of these groups is depicted here. Block arrows represent genes and are color-coded according to their predicted function/classification. Genes not

explicitly marked as pseudogenes may still be present as pseudogenes in a subset of strains within a group. For *rtgA*, versions with the start site mutation but otherwise intact are still counted as pseudogenes. The number and percentage of strains (out of 318) belonging to each group is displayed on the right.

that 615 out of 616 strains had some version of the *rtg* locus. Next, we analyzed the variation in RtgS. We found at least one copy of *rtgS* in each strain. Because duplication of *rtgS* is common, we assigned the name *rtgS1* to any copy located next to *rtgR* and the name *rtgS2* to any copy located next to *rtgR'*. Based on the identity of the penultimate residue in the translated peptide, which we have shown is important for signaling activity, we catalogued a total of three pheromone types: types A (Pro), B (Ser), and C (Leu). Only two other positions in the last 12 residues of RtgS show variation: Ala/Val at position −10 from the C-terminus and Ile/Val at position −8. The functional significance of these other polymorphisms is unknown. Finally, we analyzed each strain's RtgAB status and found 5% of strains carry unambiguously intact *rtgA* and *rtgB*. Another 12% carry intact *rtgB* and a version of *rtgA* with a start codon mutation (ATG>ATT) but is otherwise intact. We determined that a strain with the ATG>ATT mutation still produces functional RtgAB, likely by using an alternative start site, and suffers only a minor reduction in secretion capacity compared to a strain with fully intact *rtgA* (Fig. 3.9). Therefore, 17% of strains encode functional RtgAB.

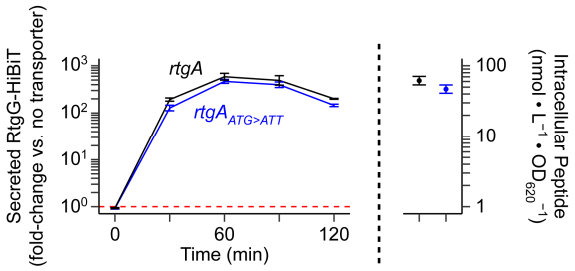

Figure 3.9. Strains with the *rtgA*$_{ATG>ATT}$ mutation still produce functional RtgAB. R6 ComAB$^-$/BlpAB$^-$/RtgAB$^-$ and RtgAB$^+$ strains with (blue) or without (black) the ATG>ATT mutation in *rtgA* expressing RtgG-HiBiT were grown in CDM+ to OD_{620} 0.05 and treated with 200 ng/mL CSP, 200 ng/mL BlpC, and 20 nM $RtgS_A$-C10. Samples were taken every 30 min and extracellular HiBiT signal was quantified (left). Data are presented as fold-change values relative to the ComAB$^-$/BlpAB$^-$/RtgAB$^-$ control. Red, dashed line represents fold-change = 1 (no difference vs. the control). At the 120-min timepoint, intracellular peptide content was also quantified (right). Plots show mean ± S.E. of 3 independent experiments.

3.3.5 Active RtgR/S confers a competitive fitness advantage during nasopharyngeal colonization.

In order to determine the biological role of *rtg*, we tested the effect of a regulatory deletion on colonization of the nasopharynx, the natural niche of pneumococcus. Despite similar levels of colonization between the wildtype and Δ*rtgR*Δ*rtgS1* strains in singly inoculated mice at 3 days post-inoculation (Fig. 3.10A), the wildtype strain outcompeted the mutant in co-inoculated mice (Fig. 3.10B). These data suggest that RtgR/S is active during nasopharyngeal colonization and show that active RtgR/S provides a fitness advantage over RtgR/S-inactive strains during co-colonization.

3.3.6 RtgAB and ComAB/BlpAB preferentially secrete different sets of peptides.

The pneumococcal PCATs ComAB and BlpAB secrete the same diverse set of GG peptides (39, 40, 206). Therefore, we wondered if ComAB and BlpAB could also secrete the *rtg* GG peptides and if RtgAB is similarly promiscuous and could secrete ComAB/BlpAB substrates. We repeated the RtgC-HiBiT and RtgG-HiBiT secretion assays with ComAB$^+$ and BlpAB$^+$ strains, using treatment with the com and blp pheromones CSP and BlpC, respectively, to induce their expression. ComAB and BlpAB secrete markedly reduced amounts of RtgC-HiBiT and RtgG-HiBiT compared to

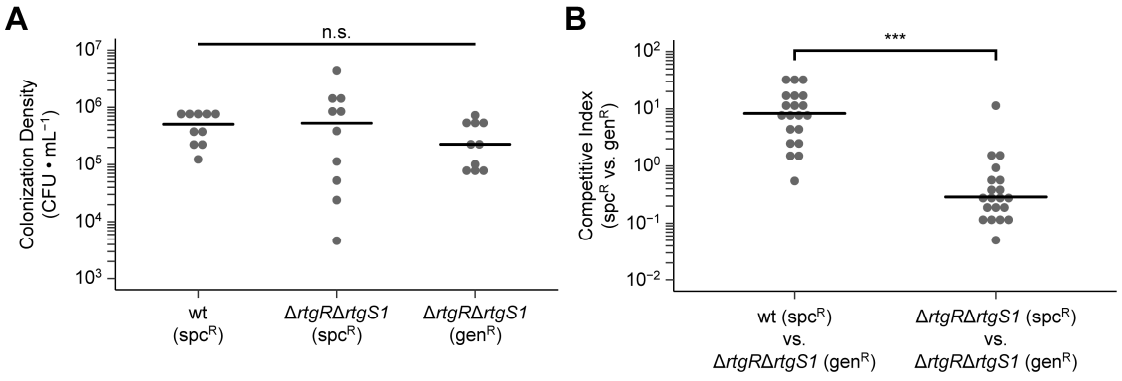

Figure 3.10. Active RtgR/S provides a competitive fitness advantage during nasopharyngeal colonization. (A) Sp9-BS68 strains were singly inoculated into the nasopharynx of 5-7 week old female BALB/c mice. At 3 days post-inoculation, colonization density was sampled by nasal wash. Black bars represent medians. N = 10 mice per strain pooled from 2 independent experiments. Statistics: n.s. not significant; Kruskal-Wallis test. **(B)** Pairs of Sp9-BS68 strains were co-inoculated into the nasopharynx of 5- to 7-week-old female BALB/c mice. At 3 days post-inoculation, colonization density was sampled by nasal wash and competitive indices were calculated. Black bars represent medians. N = 20 mice per competition pooled from 2 independent experiments. Statistics: *** $p < 0.001$; Mann-Whitney test.

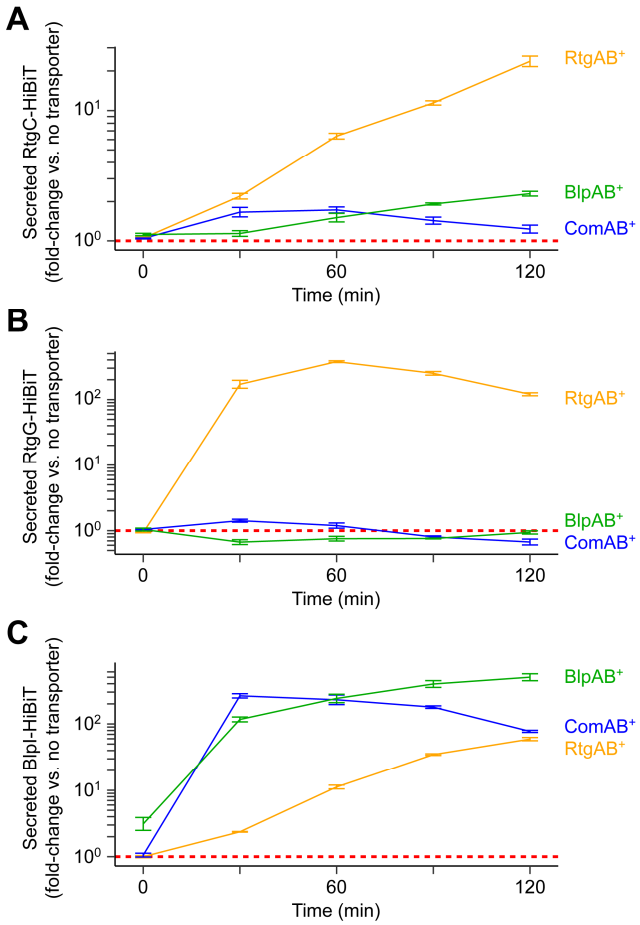

Figure 3.11. RtgAB secretes different GG peptides from ComAB and BlpAB. (A-C) R6 ComAB$^-$/BlpAB$^-$/RtgAB$^-$ and single-transporter positive ComAB$^+$, BlpAB$^+$, RtgAB$^+$ strains expressing RtgC-HiBiT (A), RtgG-HiBiT (B), or BlpI-HiBiT (C) were grown in CDM+ to OD_{620} 0.05 and treated with 200 ng/mL CSP, 200 ng/mL BlpC, and 20 nM $RtgS_A$-C10. Samples were taken every 30 min and extracellular HiBiT signal was quantified. Data are presented as fold-change values relative to the ComAB$^-$/BlpAB$^-$/RtgAB$^-$ control. Red, dashed line represents fold-change = 1 (no difference vs. the control). Plots show mean ± S.E. of 3 independent experiments.

RtgAB (Figs. 3.11A, B). To determine if RtgAB could secrete a ComAB/BlpAB substrate, we assayed secretion of a HiBiT tagged version of the BlpI bacteriocin driven by its native promoter. RtgAB secretes roughly ten-fold less BlpI-HiBiT than BlpAB (Fig. 3.11C). Therefore, RtgAB and ComAB/BlpAB do not efficiently cross-secrete each other's substrates. Consistent with this, RtgAB$^+$ strains do not show differences in *com* or *blp* activation compared to RtgAB$^-$ strains during growth in CDM+ (Fig. 3.12). Under these conditions *rtg* is expected to turn on before *com* or *blp* in every strain background.

Thus, even when RtgAB is highly expressed it secretes too little CSP and BlpC to affect the timing of *com* and *blp* activation.

3.3.7 RtgAB and ComAB/BlpAB recognize their substrates through different signal sequence motifs.

Given that we had found RtgAB and ComAB/BlpAB do not share the same substrate pool, we explored how the transporters discriminate between substrate and non-substrate GG peptides. We showed that the BlpI signal sequence (SS_{BlpI}) prevents secretion of the RtgG cargo peptide through RtgAB (Fig. 3.13A). However, it did not promote secretion of RtgG through ComAB/BlpAB, suggesting an incompatibility between the cargo peptide and these two transporters. On the other hand, the RtgG signal sequence (SS_{RtgG}) both promotes secretion of the BlpI cargo peptide through RtgAB while preventing its secretion through ComAB and BlpAB (Fig. 3.13B).

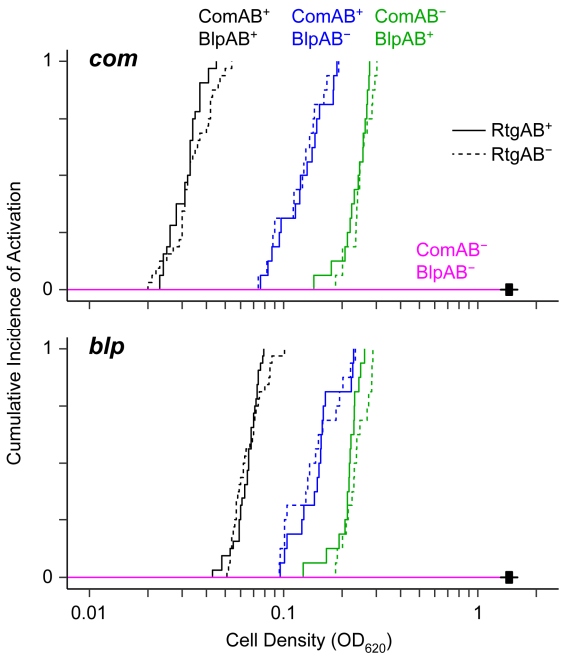

Figure 3.12. RtgAB does not affect the timing of *com* or *blp* activation during growth in CDM+. D39 P_{comA}-*Nluc* P_{BIR}-*RFluc* dual reporters were grown in CDM+ and monitored for *com* (top) and *blp* (bottom) activation. Data were fit to the Kaplan-Meier estimator. Wells that did not experience an activation event before cells reached their maximum density were censored (crosses). N = 32 ($ComAB^+$/$BlpAB^+$ strains) or 16 (all other strains) wells per strain pooled from 4 independent experiments. Statistics: all comparisons between $RtgAB^+$ and $RtgAB^-$ strains with identical ComAB/BlpAB phenotypes (except in the $ComAB^-$/$BlpAB^-$ background, for which all data points were censored) were not significant ($p \geq 0.05$); log-rank test with Holm correction.

To rule out the possibility of the secretion differences being solely caused by differences in peptide expression, we also measured the amount of intracellular peptide in each assay (Fig. 3.13A, B; right-hand graphs). The signal sequence swaps did affect intracellular peptide levels. However, these intracellular differences cannot account for the observed changes in secretion; higher intracellular levels did not correlate with more secretion, and while intracellular levels of the same peptide were relatively consistent across different strains (RtgAB$^+$ vs. ComAB$^+$ vs. BlpAB$^+$), secretion was not. Thus, the observed changes in secretion between the different peptides most likely reflect differences in peptide-transporter interactions.

In conclusion, while cargo peptide can dictate transporter compatibility in some cases, the signal sequences of GG peptides still contain all the necessary information to direct secretion of their cargo peptides through the proper transporters. For all future assays, we used BlpI as the cargo peptide since it can be secreted by all three transporters given the correct signal sequence.

Next, we searched for the specific signal sequence residues involved in transport selectivity. We found that secretion of peptide through ComAB/BlpAB depends on the identities of the residues at the conserved signal sequence positions −15, −12, −7, and −4. These positions had been previously implicated in substrate recognition by PCATs (5, 7, 8). The combination of the four residues at these positions from SS_{BlpI} (F/M/L/V) introduced into SS_{RtgG} promote secretion through ComAB/BlpAB although they are not strictly required for secretion in the context of SS_{BlpI} (Fig. 3.14A). The complementary association does not hold for RtgAB-mediated secretion in that the four residues from SS_{RtgG} (Y/L/M/L) are neither necessary nor sufficient for secretion through RtgAB (Fig. 3.14A).

In order to identify the signal sequence residues that promote secretion through RtgAB, we turned our attention to the N-terminal ends of the signal sequences, which are conserved in *rtg* GG peptides but not ComAB/BlpAB substrates. Residue swaps

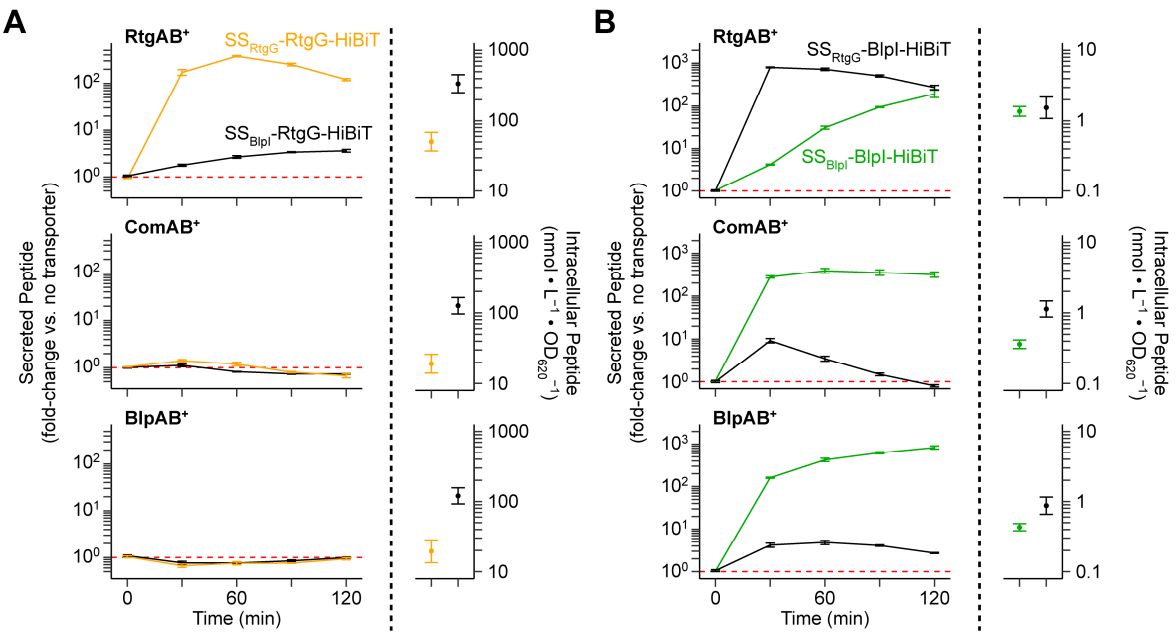

Figure 3.13. RtgAB, ComAB, and BlpAB recognize GG peptides through their N-terminal signal sequences. (A, B) R6 ComAB$^-$/BlpAB$^-$/RtgAB$^-$ and single-transporter positive ComAB$^+$, BlpAB$^+$, RtgAB$^+$ strains expressing signal sequence-swapped RtgG-HiBiT cargo peptide (A), or BlpI-HiBiT cargo peptide (B) were grown in CDM+ to OD_{620} 0.05 and treated with 200 ng/mL CSP, 200 ng/mL BlpC, and 20 nM $RtgS_A$-C10. Samples were taken every 30 min and extracellular HiBiT signal was quantified (left). Data are presented as fold-change values relative to the ComAB$^-$/BlpAB$^-$/RtgAB$^-$ control. Red, dashed line represents fold-change = 1 (no difference vs. the control). At the 120-min timepoint, intracellular peptide content was also quantified (right). Plots show mean ± S.E. of 3 independent experiments.

demonstrated that secretion through RtgAB, but not ComAB or BlpAB, depends on specific signal sequence residues at the N-terminus (positions −22 to −18) (Fig. 3.14B). These data show that unlike ComAB or BlpAB, RtgAB recognizes its substrates using a conserved motif at the N-terminal end of their signal sequences in addition to the hydrophobic residues at positions −15, −12, −7, and −4. This difference allows RtgAB to efficiently secrete a different set of peptides from ComAB and BlpAB.

3.4 Discussion

In this work, we have characterized the PCAT-encoding locus *rtg* and shown it is regulated by the RtgR/S system. RtgR/RtgS belongs to a family of regulatory systems found in streptococci that includes the Rgg/SHP and ComR/S systems (198). Rgg/SHP and ComR/S circuits can either act as cell density-dependent quorum sensing systems (198) or timing devices (213). Our data suggest RtgR/S behaves like the former (Fig.

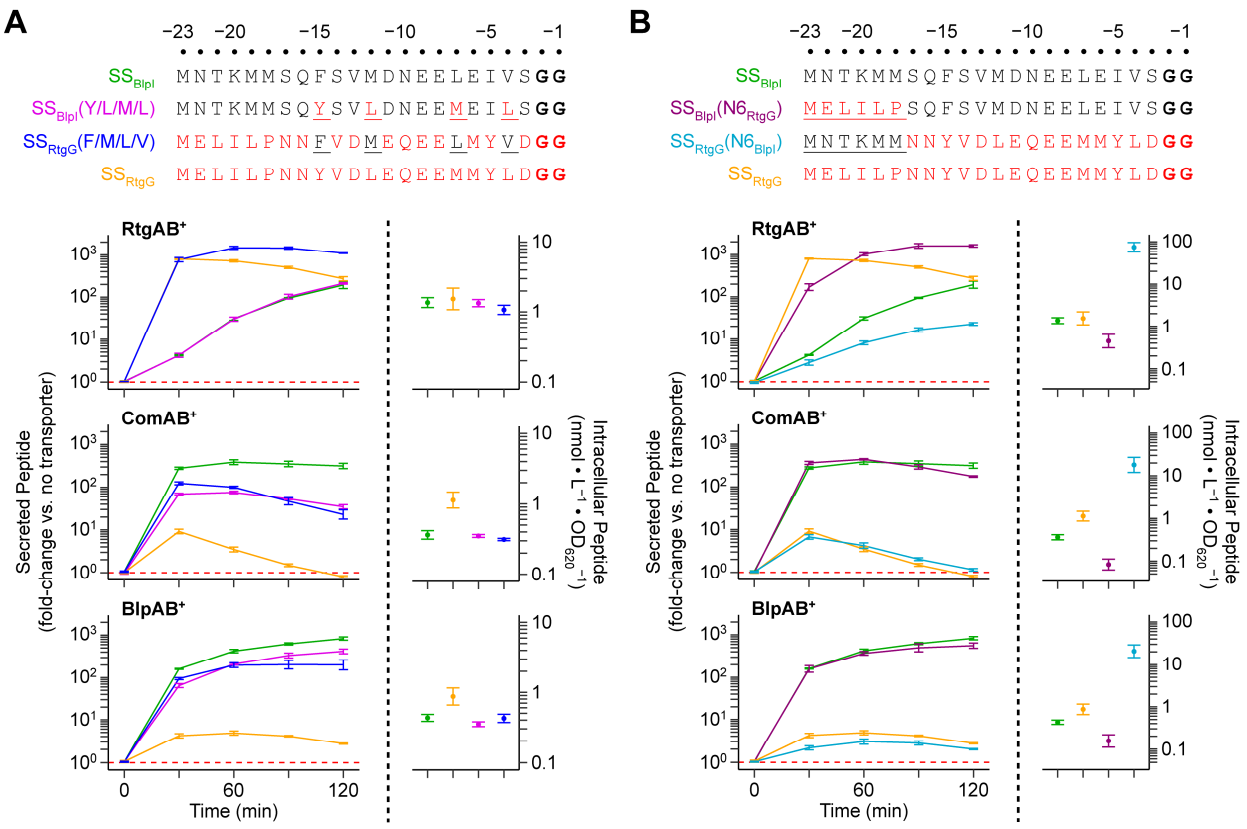

Figure 3.14. ComAB/BlpAB and RtgAB recognize substrates using different sets of signal sequence residues. (A, B) R6 ComAB$^-$/BlpAB$^-$/RtgAB$^-$ and single-transporter positive ComAB$^+$, BlpAB$^+$, RtgAB$^+$ strains expressing BlpI and RtgG signal sequences with residue swaps of the −15, −12, −7, and −4 positions (A) or N-terminal regions (B) (top; swapped residues are underlined), fused to BlpI-HiBiT cargo peptide were grown in CDM+ to OD_{620} 0.05 and treated with 200 ng/mL CSP, 200 ng/mL BlpC, and 20 nM $RtgS_A$-C10. Samples were taken every 30 min and extracellular HiBiT signal was quantified (left). Data are presented as fold-change values relative to the ComAB$^-$/BlpAB$^-$/RtgAB$^-$ control. Red, dashed line represents fold-change = 1 (no difference vs. the control). At the 120-min timepoint, intracellular peptide content was also quantified (right). Plots show mean ± S.E. of 3 independent experiments.

3.15). A purely intracellular signaling pathway has been reported for XIP in *Streptococcus mutans* (200, 214). Such a pathway is unlikely to exist for RtgR/S, since *rtg* auto-induction requires both PptAB and Ami (Fig. 3.5B). While the RtgS pheromone is similar to the previously described SHP and ComS/XIP pheromones, it also differs from these other pheromone classes in important ways. RtgS lacks the conserved aspartate or glutamate residue characteristic of SHPs and is divergently transcribed from its regulator unlike ComS (198). However, RtgS does contain a Trp-Gly-Trp motif near the C-terminus which bears resemblance to the Trp-Trp motif found in some XIPs (198, 205). RtgR is phylogenetically closer to the ComRs than SHP-associated Rgg

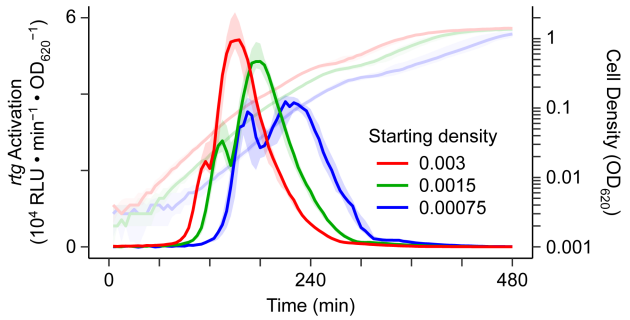

Figure 3.15. Timing of *rtg* activation depends on cell density. The wildtype Sp9-BS68 P_{rtgS1}-*luc* reporter was inoculated into CDM+ at different starting densities, grown, and monitored for *rtg* activation (dark, left y axis) and cell density (light, right y axis). Plots show median (line) and 25% to 75% quantiles (shading). N = 30 wells pooled from 3 independent experiments for each of the starting densities of 0.003 and 0.0015. N = 29 wells pooled from 3 independent experiments for starting density of 0.00075; data from one well was discarded due to a lack of growth.

regulators but does not cluster with either group (198). Using a published list of Rgg regulators (198) we found two RtgR-like regulators associated with Trp-X-Trp (WxW) motif-containing pheromones: *SPD_1518* (*rgg1518*) from *S. pneumoniae* D39 and *SSA_2251* from *Streptococcus sanguinis* SK36 (Table S1). Expression analysis of the pheromone operon associated with *rgg1518* using PneumoExpress (215) revealed that the pheromone and genes *SPD_1513* to *SPD_1517* are specifically upregulated under the same conditions that result in upregulation of *rtg*. Therefore, the Rgg1518 system is likely functional. We propose that RtgR/S and other Rgg/WxW pheromone pairs constitute a distinct group of Rgg regulatory systems. We leave the work of characterizing the members of this group and the pathways they regulate to future studies.

We showed that in the RtgAB$^+$ strain Sp9-BS68, the ability to activate the RtgR/S system confers a fitness advantage during competitive colonization of the nasopharynx. While 78% of strains are predicted to encode a functional RtgR and therefore can respond to pheromone, only 17% of strains are RtgAB$^+$. Most RtgAB$^-$ strains still encode at least one *rtg* GG peptide but have no obvious means with which to secrete them since they are not secreted by the other two PCATs commonly found in pneumococcus, ComAB and BlpAB. We have been unable to determine the function of the *rtg* GG peptides, but we speculate that they are bacteriocins. The reasons for this are that bacteriocin secretion is the most common function of PCATs, and five of the

Table 3.3. Rgg regulators associated with WxW-motif pheromones.

Species	*Strain*	*Rgg Gene*	*% Identity (% Similarity) with $RtgR_{Sp9\text{-}BS68}$*	*Predicted Pheromone Sequence*
Streptococcus pneumoniae	D39	*SPD_1518*	42.7% (62.2%)	MGFKKYLKNLPKNSGFLIWSWIQLIWFETWFWG
Streptococcus sanguinis	SK36	*SSA_2251*	45.5% (70.5%)	MKKIVYNLILLAVTSIVTTSVFPWWWLWW

seven *rtg* GG peptide genes are always associated with downstream genes encoding hypothetical proteins that resemble bacteriocin immunity proteins (216). Regardless of the specific function of the *rtg* GG peptides, the fact that most RtgAB$^-$ strains are still RtgR$^+$ suggests that *rtg* retains a useful function that does not require secretion of these peptides. Further studies will be needed to determine the mechanism responsible for the RtgR/S-dependent competitive fitness advantage seen in colonization studies, the function of the *rtg* GG peptides, and the biological significance of active *rtg* loci with non-functional RtgAB.

The case of RtgAB and ComAB/BlpAB allowed us to study how two sets of PCATs which co-exist in the same strain preferentially secrete different sets of peptides through slight differences in substrate recognition. We quantitatively assayed secretion by full-length transporters in live cells. This allowed us to potentially capture interactions and nuances that would be missed by previously employed methods that used purified, isolated PEP (6, 8, 21) or semi-quantitative measurements of secretion (36). We found that RtgAB recognizes its substrates using a motif located at the N-terminal end of their signal sequences. This motif is located 18 residues away from the signal sequence cleavage site and is exclusively found in *rtg* GG peptides. Where data are available, previous studies of PCAT substrates have found that positions at the N-terminus located farther than 18 residues from the cleavage site are either dispensable for recognition by PCATs (5, 7) or can be missing entirely (19, 38, 41). As far as we are aware, RtgAB is unique among PCATs in recognizing a signal sequence motif located so distantly from the cleavage site.

The insights into the sequence determinants of PCAT substrate selectivity gained here illuminates a relatively understudied aspect of this class of transporters. They will

also be useful in guiding future efforts to predict substrates for ComAB, BlpAB, RtgAB, and other PCATs. Some GG peptides are found without a closely associated or co-regulated PCAT (32). In these cases, it would be helpful to have sequence-based approaches to assigning potential transporters to these “orphan” GG peptides. Moreover, for strains that encode multiple PCATs, predicting if GG peptides can be secreted by PCATs that are not necessarily closely associated can guide mechanistic studies that lead to new insights into function and regulation, such as with ComAB and BlpAB substrates in pneumococcus. Our work lays the groundwork for identifying signal sequence motifs of GG peptides that are important for transporter selectivity. The next step will be to study the corresponding sequence and structural motifs in PCATs that contribute to this selectivity. In addition to bacteriocins and quorum sensing (4), GG peptides have now been linked to biofilm formation, colonization of host niches, and dissemination during infection (32, 33). Ultimately, the ability to predict and rationalize PCAT-GG peptide pairings will advance our understanding of a broad range of biologically significant microbial processes.

3.5 Materials and Methods

3.5.1 Strains and growth conditions.

All strains are derived from Sp9-BS68 (207), D39, or the R6 strain P654 (183) (referred to as PSD100 in reference) (see Table 3.4 and Appendix A for details). Pneumococcus was grown in either filter-sterilized THY (Todd Hewitt broth + 0.5% yeast extract) or CDM+ (105) at 37°C. All media contained 5 µg/mL catalase. All CDM+ was supplemented with 0.5% (v/v) THY. Except where noted otherwise, pneumococcal cultures used for experiments were inoculated to OD_{620} 0.0015 from starter cultures grown in THY pH 7.4 to OD_{620} 0.275 and frozen at −80°C in 13% glycerol. Starter cultures were pelleted at 6000×*g*, 5 min, room temperature, and resuspended in the appropriate growth media for the experiment before being used for inoculation. Antibiotics were used at the following concentrations: chloramphenicol, 2 µg/mL; gentamicin, 200 µg/mL; kanamycin, 500 µg/mL; spectinomycin, 200 µg/mL; streptomycin, 100 µg/mL.

Table 3.4. Strain list for Chapter III.

Strain	*Description*	*Antibiotic Resistances*	*Reference*
Sp9-BS68	serotype 9 clinical isolate	None	(207)
P2055	D39, *rpsL*$_{167A>C}$, Δ*blpT-blpA*::[*blpT-blpA*]$_{P32}$, ΔBIR::BIR$_{P164}$	str^R	Appendix A
P2538	R6, *rpsL*$_{167A>C}$, Δ*blpA*::*blpA*$_{P654,468\_469insAAGC}$, Δ*blpB*::*blpB*$_{P654}$, ΔBIR::BIR$_{P654}$, pE57 insertion in *blpA-blpQ* region (P$_{BIR}$-*lacZ*), Δ*comAB*, Δ*blpC*::*blpC*$_{6A}$, Δ*comC*, Δ*blpQ-pncT*::*blpI*$_{P133}$-*strepII*	cam^R, str^R	Appendix A
P2565	R6, *rpsL*$_{167A>C}$, Δ*blpA*::*blpA*$_{P654,468\_469insAAGC}$, Δ*blpB*::*blpB*$_{P654}$, ΔBIR::BIR$_{P654}$, pE57 insertion in *blpA-blpQ* region (P$_{BIR}$-*lacZ*), Δ*comAB*, Δ*blpC*::*blpC*$_{6A}$, Δ*comC*::*comC*$_{TIGR4}$, Δ*blpQ-pncT*::*blpI*$_{P133}$-*HiBiT*	cam^R, str^R	Appendix A
P2567	R6, *rpsL*$_{167A>C}$, Δ*blpA*::*blpA*$_{P654,468\_469insAAGC}$, Δ*blpB*::*blpB*$_{P654}$, ΔBIR::BIR$_{P654}$, pE57 insertion in *blpA-blpQ* region (P$_{BIR}$-*lacZ*), Δ*blpC*::*blpC*$_{6A}$, Δ*comC*::*comC*$_{TIGR4}$, Δ*blpQ-pncT*::*blpI*$_{P133}$-*HiBiT*	cam^R, str^R	Appendix A
P2569	R6, *rpsL*$_{167A>C}$, Δ*blpAB*::*blpAB*$_{P654}$, ΔBIR::BIR$_{P654}$, pE57 insertion in *blpA-blpQ* region (P$_{BIR}$-*lacZ*), Δ*comAB*, Δ*blpC*::*blpC*$_{6A}$, Δ*comC*::*comC*$_{TIGR4}$, Δ*blpQ-pncT*::*blpI*$_{P133}$-*HiBiT*	cam^R, str^R	Appendix A
P2665	D39, *rpsL*$_{167A>C}$, Δ*blpAB*::*blpAB*$_{P654}$, ΔBIR::BIR$_{P164}$, pE57 insertion in *blpA-blpI* region with Δ*lacZ*::*RFluc* (P$_{BIR}$-*RFluc*), Δ*bgaA*::P$_{comA}$-*Nluc*	cam^R, str^R	Appendix A
P2666	P2665 with Δ*blpA*::*blpA*$_{468\_469insAAGC}$	cam^R, str^R	Appendix A
P2668	P2665 with Δ*comAB*	cam^R, str^R	Appendix A
P2670	P2665 with Δ*blpA*::*blpA*$_{468\_469insAAGC}$, Δ*comAB*	cam^R, str^R	Appendix A
P2772	Sp9-BS68 with CEP-PF6-*luc*	None	Appendix A
P2775	Sp9-BS68 with CEP-P$_{rtgA}$-*luc*	None	Appendix A
P2790	P2055 with CEP-P$_{rtgS1}$-*luc*	str^R	Appendix A
P2792	Sp9-BS68 with CEP-P$_{rtgS1}$-*luc*	None	Appendix A
P2802	P2792 with Δ*rtgR*	None	Appendix A
P2804	P2792 with Δ*rtgS1*	None	Appendix A
P2811	P2792 with Δ*rtgR*, Δ*rtgS1*	None	Appendix A
P2838	P2665 with Δ*rtgAXB*::*rtgAXB*$_{Sp9-BS68}$	cam^R, str^R	Appendix A
P2840	P2666 with Δ*rtgAXB*::*rtgAXB*$_{Sp9-BS68}$	cam^R, str^R	Appendix A
P2842	P2668 with Δ*rtgAXB*::*rtgAXB*$_{Sp9-BS68}$	cam^R, str^R	Appendix A
P2844	P2670 with Δ*rtgAXB*::*rtgAXB*$_{Sp9-BS68}$	cam^R, str^R	Appendix A
P2859	P2790 with Δ*rtgS1*	str^R	Appendix A
P2878	P2538 with Δ*rtgS1*, Δ*rtgC-rtgD2*	cam^R, str^R	Appendix A
P2880	P2565 with Δ*rtgS1*, Δ*rtgC-rtgD2*	cam^R, str^R	Appendix A
P2882	P2565 with Δ*rtgS1*, Δ*rtgAXB*::*rtgAXB*$_{Sp9-BS68}$, Δ*rtgC-rtgD2*	cam^R, str^R	Appendix A
P2884	P2567 with Δ*rtgS1*, Δ*rtgC-rtgD2*	cam^R, str^R	Appendix A
P2888	P2569 with Δ*rtgS1*, Δ*rtgC-rtgD2*	cam^R, str^R	Appendix A
P2908	P2790 with Δ*rtgS2*	str^R	Appendix A
P2910	P2790 with Δ*rtgS1*Δ*rtgS2*	str^R	Appendix A
P2934	P2880 with ΔpE57, Δ*blpI*$_{P133}$-*HiBiT*::*blpQ-pncT*	str^R	Appendix A
P2936	P2882 with ΔpE57, Δ*blpI*$_{P133}$-*HiBiT*::*blpQ-pncT*	str^R	Appendix A
P2938	P2884 with ΔpE57, Δ*blpI*$_{P133}$-*HiBiT*::*blpQ-pncT*	str^R	Appendix A
P2940	P2888 with ΔpE57, Δ*blpI*$_{P133}$-*HiBiT*::*blpQ-pncT*	str^R	Appendix A
P2959	P2934 with *rtgB*-[*rtgC-HiBiT*]-*rtgD1*	str^R	Appendix A
P2961	P2936 with *rtgB*-[*rtgC-HiBiT*]-*rtgD1*	str^R	Appendix A

Strain	*Description*	*Antibiotic Resistances*	*Reference*
P2963	P2938 with *rtgB*-[*rtgC-HiBiT*]-*rtgD1*	str^R	Appendix A
P2965	P2940 with *rtgB*-[*rtgC-HiBiT*]-*rtgD1*	str^R	Appendix A
P2967	P2934 with *rtgB*-[*rtgG-HiBiT*]-*rtgH*	str^R	Appendix A
P2969	P2936 with *rtgB*-[*rtgG-HiBiT*]-*rtgH*	str^R	Appendix A
P2971	P2938 with *rtgB*-[*rtgG-HiBiT*]-*rtgH*	str^R	Appendix A
P2973	P2940 with *rtgB*-[*rtgG-HiBiT*]-*rtgH*	str^R	Appendix A
P2981	P2934 with *rtgB*-[SS_{rtgG}-*blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P2983	P2936 with *rtgB*-[SS_{rtgG}-*blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P2985	P2938 with *rtgB*-[SS_{rtgG}-*blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P2987	P2940 with *rtgB*-[SS_{rtgG}-*blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P2993	P2934 with *rtgB*-[SS_{blpI}-*rtgG-HiBiT*]-*rtgH*	str^R	Appendix A
P2995	P2936 with *rtgB*-[SS_{blpI}-*rtgG-HiBiT*]-*rtgH*	str^R	Appendix A
P2997	P2938 with *rtgB*-[SS_{blpI}-*rtgG-HiBiT*]-*rtgH*	str^R	Appendix A
P2999	P2940 with *rtgB*-[SS_{blpI}-*rtgG-HiBiT*]-*rtgH*	str^R	Appendix A
P3001	P2792 with CEP-*spcR*	spc^R	Appendix A
P3003	P2811 with CEP-*spcR*	spc^R	Appendix A
P3009	P2934 with *rtgB*-[SS_{rtgG}*(F/M/L/V)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3011	P2936 with *rtgB*-[SS_{rtgG}*(F/M/L/V)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3013	P2938 with *rtgB*-[SS_{rtgG}*(F/M/L/V)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3015	P2940 with *rtgB*-[SS_{rtgG}*(F/M/L/V)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3017	P2934 with *rtgB*-[SS_{blpI}*(Y/L/M/L)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3019	P2936 with *rtgB*-[SS_{blpI}*(Y/L/M/L)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3021	P2938 with *rtgB*-[SS_{blpI}*(Y/L/M/L)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3023	P2940 with *rtgB*-[SS_{blpI}*(Y/L/M/L)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3025	P2811 with CEP-*genR*	gen^R	Appendix A
P3027	P2934 with *rtgB*-[*blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3029	P2936 with *rtgB*-[*blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3031	P2938 with *rtgB*-[*blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3033	P2940 with *rtgB*-[*blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3035	P3001, mouse passaged	spc^R	Appendix A
P3037	P3003, mouse passaged	spc^R	Appendix A
P3039	P3025, mouse passaged	gen^R	Appendix A
P3055	P2934 with *rtgB*-[SS_{rtgG}*(N6*$_{blpI}$*)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3057	P2936 with *rtgB*-[SS_{rtgG}*(N6*$_{blpI}$*)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3059	P2938 with *rtgB*-[SS_{rtgG}*(N6*$_{blpI}$*)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3061	P2940 with *rtgB*-[SS_{rtgG}*(N6*$_{blpI}$*)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3075	P2792 with Δ*amiCD*	none	Appendix A
P3077	P2792 with Δ*pptAB*	none	Appendix A
P3081	P2934 with *rtgB*-[SS_{blpI}*(N6*$_{rtgG}$*)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3083	P2936 with *rtgB*-[SS_{blpI}*(N6*$_{rtgG}$*)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3085	P2938 with *rtgB*-[SS_{blpI}*(N6*$_{rtgG}$*)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3087	P2940 with *rtgB*-[SS_{blpI}*(N6*$_{rtgG}$*)-blpI-HiBiT*]-*rtgH*	str^R	Appendix A
P3100	P2804 with CEP-P_{rtgS1}(P2)-*luc*	none	Appendix A
P3123	P3100 with CEP-P_{rtgS1}(P2-[−82_−81TT>GG])-*luc*	none	Appendix A
P3125	P3100 with CEP-P_{rtgS1}(P2-[−80_−79GT>TG])-*luc*	none	Appendix A
P3127	P3100 with CEP-P_{rtgS1}(P2-[−65_−64GT>TG])-*luc*	none	Appendix A
P3129	P3100 with CEP-P_{rtgS1}(P2-[−46_−45AA>CC])-*luc*	none	Appendix A
P3162	P2969 with Δ*rtgA*::*rtgA*$_{ATG>ATT}$	str^R	Appendix A

3.5.2 DNA manipulation.

All PCR reactions for downstream Gibson assembly, transformation, or sequencing applications were performed using Phusion polymerase (NEB, M0530). Other PCR reactions were performed using Taq polymerase (NEB, M0273). For PCR reactions in which cells were used as template, the cells were obtained from starter cultures resuspended in sterile water and added to the PCR reaction at a 1:20 (Phusion) or 1:10 (Taq) dilution. For PCR reactions in which crude Gibson assembly product was used as template, the crude Gibson assembly product was added to the PCR reaction at a 1:100 dilution. Gibson assembly was performed using NEB HiFi DNA Assembly master mix (NEB, E2621). Primers were designed with the aid of primer3 (190, 191) and synthesized by IDT.

3.5.3 Transformations.

Transformation protocols were adapted from (187). See Appendix A for specific details on strain construction. Unmarked chromosomal mutations were created via Janus (188), Sweet Janus (189), or Janus2 (see section below) exchange. Transformants were verified by Sanger sequencing.

3.5.4 Allelic exchange using the Janus2 cassette.

The Janus2 cassette confers gentamicin resistance (200 μg/mL) and sensitivity to 10% sucrose to pneumococcal strains into which it is inserted. This allows it to be used as a counterselectable marker in the same vein as the Janus (188) and Sweet Janus cassettes (189) on which it is based. Initial selection after transforming the Janus2 cassette into pneumococcus is performed using gentamicin. Afterwards, the cassette can be exchanged with an arbitrary, markerless DNA fragment. Growth in the presence of 10% sucrose selects for transformants that have successfully exchanged the Janus2 cassette. Because the negative selection step for Janus2 does not require a streptomycin resistant copy of *rpsL* in the chromosome (or any other specific feature), Janus2 theoretically can be used in any pneumococcal strain “as is”. Also, the Janus2 cassette and the original Janus cassette (or Janus+ cassette) use orthogonal selection agents. Therefore, Janus2 can be inserted into a strain already harboring Janus, and vice versa. Exchange of the two cassettes can then be performed in either order.

Alternatively, insertion and exchange of the cassettes can each be performed in a single transformation, halving the number of transformation steps required to perform allelic exchange at two unlinked loci.

3.5.5 Sequencing of the Sp9-BS68 rtg *locus.*

To bridge a gap in the *rtg* locus in the published genome sequence of Sp9-BS68 (207), we amplified the region between *rtgG* and *rtgD2* using primers CW545/CW539 and performed Sanger sequencing using primers CW640, CW665, CW518, CW666, CW668, CW568, and CW593. This allowed us to manually and unambiguously join together contigs 306 (ABAB01000022), 342 (ABAB01000053), 99 (ABAB01000059), and 173 (ABAB01000005).

3.5.6 P_{rtgA}-luc *lytic luciferase reporter assay.*

Cells expressing P_{rtgA}-*luc* reporters were grown in THY, CDM+, or RPMI (Thermo Fisher Scientific, 11875093) ± 1% fetal bovine serum (Thermo Fisher Scientific, 10437028). At OD_{620} 0.2, cells were harvested, pelleted by centrifugation at 6000×*g*, 5 min, 4°C, washed with PBS, pelleted again and resuspended in lysis buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 1 mM DTT, 0.2% Triton X-100). After incubation with lysis buffer for 15 min at room temperature, the lysed samples were mixed with firefly luciferase reaction buffer [50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 1 mM DTT, 2 mM firefly luciferin (Thermo Fisher Scientific, 88294), 2 mM ATP, 1 mM coenzyme A] at a 1:1 ratio and luminescence was immediately read with a Synergy HTX plate reader (Biotek) in a white 96-well plate (Costar, 3917). Differences between groups were assessed by ANOVA followed by Tukey's HSD test, using the aov() and TukeyHSD() functions in R 3.5.1.

3.5.7 Luciferase reporter time course assays.

For *com*/*blp* activation assays only, starter cultures were grown in THY pH 6.8 to OD_{620} 0.075 to prevent *com*/*blp* activation. Cells were grown in THY or CDM+ in a white, clear-bottom 96-well plate (Greiner Bio-One, 655098), 200 µL per well. For assays using firefly luciferase, the following concentrations of firefly luciferin (Thermo Fisher Scientific, 88294) were added to the media: 330 µM (single reporter and dual

reporter, CDM+), 165 μM (dual reporter, THY). For assays using NanoLuc luciferase, the following concentrations of Nano-Glo substrate (Promega, N1121) were added to the media: 1:5000 (CDM+), 1:10000 (THY). The plate was incubated in a Synergy HTX plate reader, set to read absorbance at 620 nm and luminescence every 5 min. For single reporter assays, no filter was used for luminescence readings. For dual reporter assays, 450/50 band-pass and 610 long-pass filters were used to isolate NanoLuc and red firefly luciferase signals, respectively. For D39 strains only, the plate was shaken before readings were taken. Promoter activities were calculated from luminescence and absorbance readings as described in (206).

For locus activation assays, the calculations for activation event timings were adapted from ref (206). Analysis was restricted to timepoints at which OD_{620} was greater than or equal to 0.01. The low values of both luminescence and OD_{620} at timepoints before this cell density threshold resulted in low signal-to-noise ratios that made automated analysis difficult. An activation event was defined as the first timepoint at which the activation level remained above a threshold T for at least three consecutive readings. T was defined as a function of cell density as follows:

$$T = \frac{T^0}{\sqrt{N}}$$

T^0, threshold constant; N, cell density (OD_{620}). This function was chosen because it was the simplest form that empirically yielded a threshold curve with high sensitivity and specificity. The constant T^0 was empirically determined by manual inspection of activation level curves from pheromone-treated samples as positive controls and curves from non-treated pheromone deletion strains as negative controls. Wells that were already activated at the beginning of the analysis period were left censored and wells that did not activate before they reached their maximum observed cell densities were right censored.

Differences between groups were assessed by log-rank tests using the FHtest package (v1.4) in R, and when appropriate the Holm correction was applied for multiple comparisons.

3.5.8 RtgS dose-response assays.

Cells expressing P_{rtgS1}-*luc* reporters were grown in THY or CDM+ containing 330 μM firefly luciferin. At OD_{620} 0.02, cultures were aliquoted into a white, clear-bottom 96-well plate (Greiner Bio-One, 655098), 100 μL per well. Each well of the plate was prefilled with 100 μL sterile media containing 330 μM firefly luciferin and appropriate concentrations of synthetic RtgS peptide (Genscript). The plate was then incubated in a Synergy HTX plate reader set to read absorbance at 620 nm and luminescence every 5 min. For D39 strains only, the plate was shaken before readings were taken. P_{rtgS1} activity was calculated and the response was defined as the maximum observed P_{rtgS1} activity within 60 min (Sp9-BS68) or 120 min (D39) of treatment. When applicable, curves were fit to a Hill model using the nls() function in R 3.5.1.

3.5.9 Peptide secretion assays.

Cells were inoculated from starter cultures to OD_{620} 0.005 and grown in CDM+. At OD_{620} 0.05, cells were treated with 200 ng/mL CSP1, 200 ng/mL $BlpC_{R6}$, and 20 nM $RtgS_A$-C10. Samples were taken for HiBiT quantification at appropriate time points. For native BlpI-HiBiT assays only, clarified supernatants were obtained after centrifugation at 6000×*g*, 5 min, 4°C. For all other assays, cells were retained in the samples. HiBiT signal was quantified by mixing samples with HiBiT Extracellular Detection Reagent (Promega, N2421) at a 1:1 ratio in a white 96-well plate (Costar, 3917) and reading luminescence with a Synergy HTX plate reader after a 1-min shake step. Samples were also taken for quantification of intracellular peptide; for endpoint assays, they were taken concurrently with the extracellular samples, and for time-course assays, they were taken at the last time point. These samples were pelleted at 6000×*g*, 5 min, 4°C and resuspended in one-fifth original volume proteinase K buffer [20 mM MES pH 6.5, 20 mM $MgCl_2$, 0.5 M sucrose, 100 μg/mL proteinase K (Fisher Scientific, BP1700)]. A 15-min incubation at 37°C removed residual extracellular peptide by proteinase K digestion. Afterwards, proteinase K was inactivated by addition of 1 mM phenylmethanesulfonyl fluoride (Calbiochem, 7110) and cells were lysed by mixing with 2% Triton X-100 at a 1:1 ratio followed by incubation at room temperature for 15 min. The lysed samples were then mixed with HiBiT Extracellular Detection Reagent at a 1:1

ratio in a white 96-well plate and read with a Synergy HTX plate reader as above. Standards consisting of synthetic L10-HiBiT peptide (206) mixed with lysates of a non-HiBiT expressing strain were used to generate standard curves to use for calculating HiBiT-tagged peptide concentrations in experimental samples. Differences between groups were assessed by ANOVA followed by Tukey's HSD test, using the aov() and TukeyHSD() functions in R 3.5.1.

3.5.10 Genomic analysis of rtg*.*

Genomes of the Massachusetts pneumococcal isolate collection (BioProject Accession: PRJEB2632) were filtered based on sequence coverage of the predicted location of *rtg*. Three possible upstream flanking genes were defined: a gene coding for a LysM-domain protein (*SPD_0104*), *argG* (*SPD_0110*), and *argH* (*SPD_0111*). Two possible downstream flanking genes were defined: a putative endoRNAse gene (*SPD_0125*; disrupted in D39) and *pspA* (*SPD_0126*). The D39 versions of these genes were BLASTed (megablast profile) against all genomes in the Massachusetts collection. In the case of *pspA*, only the first 250 nucleotides were used so that the choline-binding repeats found later in the gene did not complicate the analysis. All genomes in which at least one upstream flanking gene was found on the same contig as at least one downstream flanking gene were chosen for further analysis (318 in total); all other genomes were discarded. For each genome in the filtered set of 318, the sequence between the two closest flanking genes was extracted and a distance matrix for these sequences representing *rtg* loci was calculated using BIGSdb (217). The set of unique *rtg* genes found in Sp9-BS68 and D39 was used as the reference gene list. Then, a neighbor-joining tree was constructed from the distance matrix using SplitsTree (218). This tree was used to guide manual clustering of the *rtg* loci into different groups based on gene presence and synteny. New unique *rtg* genes found in the course of this analysis were added to the initial set of genes from Sp9-BS68 and D39 and the updated set was BLASTed (blastn profile) against all 616 genomes in the Massachusetts collection to determine how many strains in the full collection encoded *rtg*. Genomes with a hit (≥ 80% query coverage, ≥ 70% sequence identity) for at least one *rtg* gene were considered to be *rtg*-positive.

3.5.11 Mouse colonization assays.

Mouse colonization was performed as described in (206). Briefly, dual or single-strain mixtures of Sp9-BS68 were inoculated into the nasopharynx of un-anaesthetized 5- to 7-week-old female BALB/c mice (Taconic). Mice were euthanized with CO_2 overdose after 72 hours and nasopharyngeal colonization was sampled by nasal wash.

Ten-fold serial dilutions of PBS suspensions used to inoculate mice were spot plated (5 µL per spot, 3 replicates per sample) on tryptic soy agar (TSA) plates supplemented with 5 µg/mL catalase and containing either neomycin (selects for all pneumococcus), spectinomycin, or gentamicin. Plates were incubated at 37°C, 5% CO_2 overnight. Afterwards, colony counts were obtained. To obtain colonization density data from singly inoculated mice, nasal washes were quantified in the same manner as described above, and counts from the neomycin plates were used. When attempting to quantify colonization density from mice inoculated with two competing strains, we noticed that the growth of the gentamicin-resistant strain on gentamicin-containing plates were inhibited by an unknown factor from the nasal wash. This occurred during both trials of the colonization assay. This inhibition was not seen for the gentamicin-resistant strain growing on neomycin-containing plates from the singly inoculated mice. No inhibition was seen for either of the spectinomycin-resistant strains growing on either spectinomycin- or neomycin-containing plates. This inhibition was also not seen for any strain growing on any plate during plating of the inoculants.

The above phenomenon prevented us from obtaining accurate counts of the gentamicin-resistant strain from co-inoculated mice using the spot-plating method. However, since the inhibition was restricted to only gentamicin-containing plates, we developed an alternative method for quantifying the ratio of spectinomycin-resistant to gentamicin-resistant colonies isolated from co-inoculated mice using the samples plated on neomycin-containing plates. For each mouse, we picked 40-100 colonies from the spots on the neomycin-containing plates, grew them in THY + 5 µg/mL catalase at 37°C for 6-8 hours, then replicate-plated the cultures on TSA plates supplemented with 5 µg/mL catalase and containing either spectinomycin or gentamicin. After an overnight incubation at 37°C with 5% CO_2, each culture was scored for growth on spectinomycin or gentamicin. Cultures that showed growth on both antibiotics were excluded from

further analysis. If none of the cultures from a mouse grew on a particular antibiotic, these data points were treated as being below the limit of detection and a count of 1 was used instead of 0. For each mouse, the output ratio was defined to be the count of cultures showing spectinomycin-resistant growth divided by the count of cultures showing gentamicin-resistant growth. We confirmed that neither of the spectinomycin-resistant strains could inhibit growth of the gentamicin-resistant strain or vice versa during growth on neomycin-containing plates or in co-culture in THY by repeating the above method on spot-plated samples from mixed cultures prepared in an identical fashion to those used to inoculate mice. Ratios obtained from the replicate-plating method agreed closely with ratios obtained from traditional counting of colonies on spectinomycin- and gentamicin-containing plates. The final competitive index calculation was performed using the following formula:

$$CI = R_{out}\frac{N_{gen,in}}{N_{spc,in}}$$

CI, competitive index; R_{out}, output ratio; $N_{gen,in}$, input CFU density of gentamicin-resistant strain; $N_{spc,in}$, input CFU density of spectinomycin-resistant strain.

Differences in colonization densities and competitive indices between groups were evaluated by the Mann-Whitney (2 groups) and Kruskal-Wallis (>2 groups) tests using the wilcox.test() and kruskal.test() functions in R 3.5.1.

Chapter IV: Conclusions

In my thesis, I studied three pneumococcal transporters of the PCAT family: ComAB, BlpAB, and RtgAB. In chapter II, I examined the degree of substrate sharing between ComAB and BlpAB and how peptide secretion by these transporters drive functional differences related to competence and pneumocins between BlpAB$^+$ and BlpAB$^-$ strains. In chapter III, I characterized the *rtg* locus and the PCAT encoded within, RtgAB, and showed how RtgAB and ComAB/BlpAB maintain separate substrate pools.

I determined that ComAB and BlpAB secrete the same set of GG peptides, including the pheromones CSP and BlpC and the pneumocins. Cross-secretion of the pheromones mediates bidirectional crosstalk between the *com* and *blp* regulatory systems, allowing them to temporally coordinate their activation. It has been shown that *com*-to-*blp* crosstalk facilitates DNA exchange during competence (40). Crosstalk in the opposite direction may be useful because it allows cells to activate competence in a timely manner to exploit the reservoir of DNA liberated by pneumocin-mediated killing. This hypothesis was not tested here but should be a priority for future studies. Additionally, secretion of CSP by BlpAB even in the absence of *blp* activation increases the propensity with which competence activation occurs in BlpAB$^+$ strains compared to BlpAB$^-$ strains. This would allow BlpAB$^+$ strains to activate competence more frequently, at lower cell densities, and under a wider range of conditions. Follow-up studies will be needed to determine how this affects the evolution and fitness of BlpAB$^+$ and BlpAB$^-$ strains.

The discovery that ComAB can secrete the pneumocins provides an explanation for why certain strains encode pneumocins but not functional BlpAB; these strains secrete pneumocins using ComAB instead. A previous study had reported differences in the capacity of BlpAB$^+$ and BlpAB$^-$ strains to produce pneumocin-mediated killing (14). I showed that this can be explained by differences in pneumocin regulation between the two populations and differences in ComAB and BlpAB regulation. BlpAB$^-$ strains can only produce pneumocins during competence activation while BlpAB$^+$ strains are not as restricted. Additionally, BlpAB secretes greater total amounts of pneumocin during *blp* activation than ComAB. This is because ComAB expression when activated is limited in

duration by the competence shut-off mechanism whereas BlpAB expression is not. Therefore, BlpAB$^+$ strains can induce pneumocin expression more frequently and secrete more pneumocins when induced than BlpAB$^-$ strains. This is corroborated by in vivo data showing that BlpAB$^+$ strains directly outcompete BlpAB$^-$ strains in a mouse model of nasopharyngeal colonization. With respect to pneumocin use, these data show BlpAB$^+$ strains are the aggressors, effective at deploying pneumocins to kill competitors under general conditions. Meanwhile, BlpAB$^-$ strains are the opportunists, using pneumocins in a more limited manner, presumably to support competence. The fact that a 25% BlpAB$^+$/75% BlpAB$^-$ split has been maintained throughout the pneumococcal population suggests that both the aggressor and opportunist phenotypes are useful in different contexts. The types of contexts that favor one phenotype over the other are yet to be defined. Moreover, further inquiries into how pneumocin-mediated killing shapes colonization dynamics in humans are needed in order to understand how BlpAB$^+$ and BlpAB$^-$ strains interact with each other in their natural hosts.

In addition to ComAB and BlpAB, I characterized the substrate range of RtgAB, a PCAT encoded in the *rtg* locus. RtgAB and the rest of the locus are regulated by the Rgg/SHP-like RtgR/S system, which confers a competitive fitness advantage during nasopharyngeal colonization in mice. RtgAB secretes a set of co-regulated *rtg* GG peptides and secretes *com*- and *blp*-regulated GG peptides only at low levels. Neither ComAB nor BlpAB secretes the *rtg* GG peptides. I investigated the sequence determinants of substrate selectivity between these two groups of transporters and found that selectivity is determined by specific GG peptide signal sequence motifs. ComAB and BlpAB prefer substrates with specific combinations of hydrophobic residues at four conserved signal sequence positions previously reported to be important for substrate recognition (8). On the other hand, RtgAB tolerates a wider range of residues at these positions but requires for efficient secretion a motif exclusively found in the *rtg* GG peptides located at the N-terminal end of their signal sequences. ComAB and BlpAB have no such requirement for an N-terminally situated motif. These data show how two sets of pneumococcal PCATs have evolved subtly different ways of recognizing their substrates in order to preferentially secrete different sets of peptides.

Many PCATs have been reported to be able to secrete heterologous substrates (7, 21, 22, 38). However, there is comparatively little data on how PCATs found in the same strain of bacteria distinguish between different GG peptides. The case of ComAB and BlpAB in pneumococcus shows how substrate sharing between PCATs can have significant functional implications on the pathways in which they participate. Therefore, when PCATs have evolved to be selective, the details of how they recognize substrate, but not non-substrate GG peptides are an important part of their biology worthy of study. Additionally, being able to predict if a PCAT can secrete a particular GG peptide in the absence of experimental data would facilitate the study of PCAT/GG peptide-dependent processes. Such predictive abilities would also aid in the search for the specific exporters of GG peptides found without an obvious associated PCAT, such as VP1 from pneumococcus (32). The work on the sequence determinants of GG peptide export through RtgAB and ComAB/BlpAB lays the groundwork for associating certain classes of GG peptide signal sequences with their corresponding PCATs. If the distinguishing signal sequence motifs of GG peptides are the keys with which they gain access to specific PCATs, then the next step will be to identify the sequence and structural elements of PCATs that form the locks which recognize the keys. Given that the peptidase domains of PCATs have been shown to bind the signal sequences of GG peptides, these lock elements will most likely be found there.

There are several unanswered questions regarding RtgAB and the *rtg* locus in general. First, the mechanism underlying the competitive fitness advantage conferred by RtgR/S during nasopharyngeal colonization has not been determined. The mechanism is likely associated with the rest of the *rtg* locus, though I have not ruled out the possibility of RtgR/S regulating another, ectopic locus that is responsible for the observed fitness advantage. Second, the specific function of the *rtg* GG peptides is still unknown. Currently, the favored hypothesis is that these peptides are bacteriocins. The most common roles of PCATs are to secrete bacteriocins and signaling peptides. There are seven distinct *rtg* GG peptides (eight if the fusion protein RtgV is included), of which between zero and five may be encoded by a single strain. It would be unusual for a locus to encode so many different signaling peptides and for different strains to possess different numbers of them. On the other hand, bacteriocin loci often include multiple

bacteriocin peptides (15, 18, 138). The diversity in *rtg* GG peptide content across different strains is similar to the variation observed at the *blp* bacteriocin locus (138). Additionally, five of the *rtg* GG peptide genes (*rtgC*, *rtgG*, *rtgK*, *rtgM*, *rtgT*) are always found associated with small genes located immediately downstream (*rtgD1*, *rtgH*, *rtgL*, *rtgN*, *rtgU*, respectively) which code for hypothetical proteins resembling bacteriocin immunity proteins (216). So far, I have been unable to experimentally demonstrate antimicrobial activity attributable to the *rtg* locus using standard overlay assays in agar plates (219). However, there are technical difficulties associated with inducing expression of *rtg* in agar plates that could be responsible for the absence of observed antimicrobial activity. Efforts should be made in the future to modify the overlay assay or develop a new one better suited to studying the potential antimicrobial activity of the *rtg* GG peptides.

Another unanswered question involves the functional significance of *rtg* loci with disrupted *rtgAB* genes. Strains encoding such loci have no obvious method of secreting the *rtg* GG peptides. Yet the majority of these RtgAB$^-$ strains still encode an intact RtgR regulator which can induce *rtg* in response to either self- or exogenously produced RtgS pheromone. Since expending resources to activate a locus that has no use is presumably evolutionarily disfavored, this suggests that *rtg* retains some function even in the absence of RtgAB. If the *rtg* GG peptides are in fact bacteriocins, then production of immunity proteins would still be advantageous for RtgAB$^-$ strains to protect themselves against bacteriocins produced by competing strains since the transporter is not required for immunity protein maturation. This would be analogous to the “cheater” hypothesis previously proposed to explain why the majority of strains encode an inducible *blp* locus but non-functional BlpAB (14). Since RtgAB$^-$ strains make up 83% of the pneumococcal population, determining the role that *rtg* plays in these strains should be a priority for follow-up studies.

Besides providing insights into the specific systems under study, the work presented in this thesis also contributes to our general understanding of how PCATs secrete GG peptides. This family of peptides is associated with a diverse assortment of pathways such as bacteriocin-mediated killing (4), quorum sensing (24, 28, 196), dissemination during infection (32), and biofilm formation (33). Therefore, the

knowledge gained in this area carries broad significance for the fields of microbial physiology and pathogenesis.

APPENDIX A: Construction of Strains

Construction of the R6 com/blp *dual luciferase reporter.*

To create the P_{comAB}-Nluc reporter, first the NanoLuc gene (Promega) was codon optimized for *S. pneumoniae* R6. This optimized sequence was then placed immediately downstream of the *S. pneumoniae rpsD* terminator followed by the *comAB* promoter from R6 (nucleotides −401 to −1 relative to the start triplet of *comA*). This fragment (1xterm-P_{comAB}-*Nluc*) was synthesized by IDT using their gBlocks service and subsequently PCR amplified using primers CW117/CW194. For double-crossover integration into the *bgaA* locus, upstream and downstream fragments were created as follows. The upstream fragment (bgaA-up) was PCR amplified from P654 using primers CW121/CW115. The downstream fragment (3xterm-bgaA-down) was PCR amplified from a P654 derivative carrying a *lacZ* reporter integration in *bgaA* (replacing nucleotides 91-6582) followed by the *B. subtilis rrnB* and *S. pneumoniae rpsI* and *tufA* terminators using primers CW183/CW122. The sequences of the transcriptional terminators were obtained from ref (220). The three fragments, bgaA-up, 1xterm-P_{comAB}-*Nluc*, and 3xterm-bgaA-down, were ligated together using Gibson assembly and the crude product was PCR amplified using primers CW109/CW112 to create the final assembled product, *bgaA*-P_{comAB}-*Nluc*. This product was transformed into pneumococcus via Janus exchange (188) as follows. A PCR product containing the Janus cassette inserted into *bgaA* was amplified using primers CW109/CW112 following Gibson assembly of the following PCR products: 1) primers CW121/CW110 from P654, 2) primers CW105/CW106 from a pneumococcal strain containing the Janus cassette, 3) primers CW111/CW122 from P654. This product was used to transform P654 to create strain P2163. P2163 was then transformed with bgaA-P_{comAB}-*Nluc* to create strain P2190.

To create the P_{BIR}-RFluc reporter, first the red firefly luciferase gene (Targeting Systems) was codon optimized for *S. pneumoniae* R6 to create the *RFluc* gene. *RFluc*

was synthesized by Genscript and PCR amplified using primers CW199/CW200. For double-crossover integration to replace *lacZ* in the integrated pE57 plasmid in P654, upstream and downstream fragments were created as follows. The upstream fragment (BIR-up) was amplified from P654 using primers CW39/CW201 and the downstream fragment (*lacZ*-down) was amplified from P654 using primers CW202/CW83. The three fragments, BIR-up, *RFluc*, and *lacZ*-down, were ligated together using Gibson assembly and the crude product was PCR amplified using primers CW81/CW82 to create P_{BIR}-*RFluc*. This product was transformed into pneumococcus via Janus exchange as follows. A PCR product containing the Janus cassette inserted into *lacZ* was amplified using primers CW81/CW82 from a strain containing the Janus cassette inserted between the two HpaI restriction sites in *lacZ*. This PCR product was used to transform P2190 to create strain P2208. P2208 was then transformed with P_{BIR}-*RFluc* to create strain P2219, the wild-type R6 dual luciferase reporter.

Construction of the D39 com/blp *dual luciferase reporter.*

A streptomycin-resistant D39 strain ($rpsL_{167A>C}$) was transformed to replace the entire *blp* locus via Janus exchange with crude lysate from P213 (3) to create P2055. P2055 was then transformed with crude lysate from P2219 followed by chloramphenicol selection, and crude lysate from the resulting transformant was used to back-transform P2055 to create P2610. P2610 has the native D39 alleles of *blpT* through *blpC*, the P654 allele of *blpB* (originally derived from a serotype 6B strain), and the P654 allele of *blpA* (originally derived from a serotype 6B strain) with two substitutions: c.1472A>G and c.1692C>A. The BIR of P2610 contains the following genes (in order): *blpI*, *blpU4*, *blpU5*, *blpJ*, *blpW1*, *blpW2*, *blpG*, *blpL'*, *tdpA*, *hypothetical protein* (273 bp), *hypothetical protein* (357 bp), *pncG*, *blpL*, *blpZ*, *pncP*.

P2610 was then transformed with the *bgaA*-P_{comAB}-*Nluc* PCR product via Sweet Janus exchange (4) to create strain P2625. The Sweet Janus-containing PCR product used in this transformation was obtained as follows. A DNA fragment containing the Sweet Janus cassette was synthesized by IDT using their gBlocks service and amplified using primers CW105/CW106. Upstream and downstream homology fragments were PCR amplified from P654 using primers CW121/110 and CW111/122. These three

fragments were ligated together using Gibson assembly and the crude product was PCR amplified using primers CW109/112.

*Construction of the P_{BIR}-*luc* reporter.*

First, the pE57 plasmid insertion in P654 was excised by passaging in chloramphenicol(−) media to create strain P1613. Next, the *Photinus pyralis luc* gene was PCR amplified from a pneumococcal strain containing a pR422 insertion (5) using primers SD1/SD2, and the resulting product was digested with BamHI/StuI (NEB) and ligated into linearized pE57 cut with BamHI/EcoRV (NEB). The resulting plasmid, pE135, was then transformed into P1613 to create the wild-type reporter strain P1666, which is effectively isogenic to P654 except that it has a P_{BIR}-*luc* construct in place of a P_{BIR}-*lacZ* construct.

Construction of the Sweet Janus+ cassette.

The Sweet Janus cassette was modified as follows to increase kanamycin resistance and reduce the rate of spontaneous reversion to a streptomycin resistant phenotype. First, a DNA fragment was synthesized by IDT using their gBlocks service containing the 3' end of *aphA-3* immediately followed by the promoter region of the *rplM* gene from R6 transcriptionally fused to a version of *rpsL+* with five silent substitutions with respect to the wild-type R6 allele: c.111C>T, c.123A>G, c.150T>C, c.366T>C, c.405A>G. This fragment was PCR amplified using primers CW358/CW106. Next, to replace the *sacB-aphA3* intergenic region of the Sweet Janus cassette with a modified version containing an optimized ribosome binding site, two DNA fragments were amplified from a pneumococcal strain containing the Sweet Janus cassette using primers CW105/CW375 and CW376/CW357. All three DNA fragments were then ligated together using Gibson assembly to create the Sweet Janus+ cassette. The optimized ribosome site upstream of *aphA3* increases kanamycin resistance. The silent substitutions in *rpsL+* decrease the rate of spontaneous Str^R reversion by *hex*-dependent recombination inhibition (188). The inclusion of the dedicated *rplM* promoter to drive high-level *rpsL+* expression also reduces spontaneous Str^R reversion through an unknown mechanism. The nucleotide sequence of the Sweet Janus+ cassette is deposited in GenBank with accession number MH304214.

Construction of unmarked transporter and pheromone mutants.

An unmarked Δ*comAB* mutation that deletes the entire *comAB* operon was created by Gibson assembly of two fragments obtained from PCR of P654 using primers CW33/CW34 and CW35/CW36, followed by PCR of the crude assembly product using primers CW33/CW36. This PCR product was used to transform R6 strains via Janus exchange and D39 strains via Sweet Janus+ exchange. The Janus-containing PCR product (*comAB*-Janus) used in these transformations was obtained by amplification of P1537 (40) using primers CW33/CW36. The Sweet Janus+ -containing PCR product used in these transformations was obtained as follows. The Sweet Janus+ cassette was assembled as described above. Upstream and downstream homology fragments were PCR amplified from P654 using primers CW268/CW182 and CW359/CW92. These three fragments were ligated together using Gibson assembly and the crude product was PCR amplified using primers CW84/CW360.

An unmarked Δ*blpA* mutation that introduces a 4-bp frameshift duplication (c.468_469insAAGC) was created by Gibson assembly of two fragments. Fragment one was obtained from PCR of either P654 using primers CW102/CW131, P1666 using primers CW100/CW131, or P2219 using primers CW242/CW131. Fragment two was obtained from PCR of P654 using primers CW132/CW167. The final product was PCR amplified from the crude assembly product using either primers CW150/CW46 (P654-derived), CW166/CW46 (P1666-derived), or CW241/CW46 (P2219- derived) to create P654-$blpA_{fs}$, P1666-$blpA_{fs}$, or P2219-$blpA_{fs}$, respectively. These PCR products were used to transform pneumococcal strains containing P_{BIR}-*lacZ*, P_{BIR}-*luc* or P_{BIR}-*RFluc* reporters via Janus exchange for R6 strains or Sweet Janus+ exchange for D39 strains. The Janus-containing PCR product used in these transformations was obtained as follows. Fragment one was obtained from PCR of either P654 using primers CW102/CW58 or P2219 using primers CW242/CW58. Fragment two was obtained from PCR of a Janus-containing strain using primers CW59/CW60. Fragment three was obtained from PCR of P654 using primers CW61/CW167. These three fragments were ligated together using Gibson assembly, and the final product was PCR amplified from the crude Gibson assembly product using either primers CW150/CW46 (P654-derived) or CW241/CW46 (P2219-derived). The Sweet Janus+ -containing PCR product used in

these transformations was obtained as follows. The Sweet Janus+ cassette assembled as described above. Upstream and downstream homology fragments were PCR amplified from P2625 using primers CW242/CW58 and CW61/CW167. These three fragments were ligated together using Gibson assembly and the crude product was PCR amplified using primers CW241/CW46.

An unmarked Δ*comC* mutation that deletes nucleotides +31 to +123 of *comC* was created by Gibson assembly of two fragments obtained from PCR of P654 using primers CW13/CW20 and CW19/CW24, followed by PCR of the crude assembly product using primers CW13/CW24. This PCR product was used to transform pneumococcus via Janus exchange. The Janus-containing PCR product used in these transformations (*comC*-Janus) was obtained by Gibson assembly of three fragments obtained from PCR of P654 using primers CW13/CW14 and CW23/CW24 and of a Janus-containing strain using primers CW15/CW16, followed by PCR of the crude assembly product using primers CW13/CW24.

An unmarked Δ*blpC* mutation that deletes nucleotides +28 to +129 of *blpC* was created by Gibson assembly of two fragments obtained from PCR of P654 using primers CW45/CW246 and CW247/CW78, followed by PCR of the crude assembly product using primers CW74/CW77. This PCR product was used to transform pneumococcus via Janus exchange. The Janus-containing PCR product used in these transformations (*blpC*-Janus) was obtained by Gibson assembly of three fragments obtained from PCR of P654 using primers CW45/CW245 and CW105/CW106 and of a Janus-containing strain using primers CW108/CW78, followed by PCR of the crude assembly product using primers CW74/CW77.

Construction of mismatched com *pheromone/receptor strains.*

The CSP2-ComD1 mismatch was created in R6 as follows. Sequence of the type 2 *comC* allele from TIGR4 was synthesized by IDT using their gBlocks service and PCR amplified using primers CW171/CW170. Upstream and downstream homology fragments were PCR amplified from P654 using primers CW147/CW172 and CW173/CW148. These three fragments were ligated together using Gibson assembly and the crude product was amplified using primers CW13/CW24. This final product was

used to transform pneumococcus via Janus exchange with the *comC*-Janus PCR product.

Construction of mismatched blp *pheromone/receptor strains.*

The $BlpC_{6A}$-$BlpH_{R6}$ mismatch was created in R6 as follows. Sequence of the type 6A *blpC* allele (135) was synthesized by IDT using their gBlocks service and PCR amplified using primers CW144/CW80. Upstream and downstream homology fragments were PCR amplified from P654 using primers CW45/CW145 and CW143/CW78. These three fragments were ligated together using Gibson assembly and the crude product was amplified using primers CW74/CW77. This final product was used to transform pneumococcus via Janus exchange with the *blpC*-Janus PCR product.

The $BlpC_{R6}$-$BlpH_{6A}$ mismatch was created in R6 as follows. First, a PCR product containing a replacement of *blpH* to *blpC* with the Janus cassette was created. Two fragments were PCR amplified from a strain transformed with the *blpC*-Janus PCR product with primers CW255/CW257 and CW106/CW165. These fragments were ligated together using Gibson assembly and the crude product was amplified using primers CW256/CW45 to create *blpHC*-Janus. Next, sequence of a *blpH* allele encoding a type R6/6A chimeric BlpH that is responsive to type 6A pheromone was PCR amplified from P1039 (183) (referred to as PSD118 in reference) using primers CW266/CW143. Upstream and downstream homology fragments were PCR amplified from P654 using primers CW255/CW265 and CW80/CW165. These three fragments were ligated together using Gibson assembly and the crude product was amplified using primers CW256/CW45. This final product was used to transform pneumococcus via Janus exchange with the *blpHC*-Janus PCR product.

Construction of strains expressing tagged ComC and BlpI.

R6 strains expressing ComC-FLAG were created as follows. First, a DNA fragment containing *comC* with a codon-optimized 2xFLAG sequence (gattacaaagatgacgatgataaagactacaaagacgacgatgacaaa) inserted between the last coding triplet and the stop triplet was synthesized by IDT using their gBlocks service and subsequently PCR amplified using primers CW171/CW170. Next, upstream and downstream homology fragments were PCR amplified from P654 using primers

CW147/CW172 and CW173/CW148. These three fragments were ligated together using Gibson assembly and the crude product was amplified using primers CW13/CW24 to create *comC-FLAG*. P654 was transformed as described above to create transporter deletion mutants, then *comC-FLAG* was used to transform P654 or these mutants via Janus exchange with the *comC*-Janus PCR product to create strains P2141 (wild-type), P2211 (Δ*comAB*), P2213 (Δ*blpA*), and P2215 (Δ*comAB*Δ*blpA*).

R6 strains expressing BlpI-FLAG were created as follows. First, a sequence containing a replacement of *blpQ* to *pncT* with *blpI* from strain P133 (14) carrying an insertion of a codon-optimized 2xFLAG sequence (see above) inserted between the last coding triplet and the stop triplet, along with roughly 1 kb of flanking P654 sequence on each end was synthesized by Genscript. This sequence was PCR amplified using primers CW213/CW219 and used to transform P654 via Janus exchange to create strain P2309. The Janus-containing PCR product used in these transformations (*blpQ-pncS*-Janus) was obtained by Gibson assembly of three fragments obtained from PCR of P654 using primers CW212/CW232 and CW238/CW234, and of a Janus-containing strain using primers CW105/CW106, followed by PCR of the crude assembly product using primers CW213/CW219. Next, P2309 was transformed as described above to express mismatched *blp* type 6A pheromone/type R6 receptor, creating strain P2329. Finally, P2329 was transformed as described above to create the transporter deletion mutants P2341 (Δ*blpA*), P2343 (Δ*comAB*), and P2363 (Δ*comAB*Δ*blpA*).

R6 strains expressing BlpI-HiBiT were created as follows. First, a DNA fragment containing an insertion of a codon-optimized HiBiT sequence preceded by a -GGGGSGGGGS- linker (ggtggtggaggttcaggaggtggaggttctgtttctggttggcgtctttttaaaaaaatttca) between the last coding triplet and the stop triplet of $blpI_{P133}$ was synthesized by IDT using their gBlocks service. This fragment was PCR amplified using primers CW262/CW263. Next, upstream and downstream homology fragments were PCR amplified from P2329 using primers CW212/CW261 And CW264/CW234. These three fragments were ligated together using Gibson assembly and the crude product was amplified using primers CW213/CW219. This final product was used to transform P2329, P2341, P2343, and P2363 via Janus exchange with the *blpQ-pncS*-Janus PCR product to create strains

P2475, P2483, P2485, and P2477, respectively. To create strains expressing mismatched *com* type 2 pheromone/type 1 receptor, P2475, P2483, P2485, and P2477 were transformed as described above to create strains P2563, P2569, P2567, P2565, respectively. The Δ*comAB*Δ*blpA* background control strains, P2384 (wild-type *comC*) and P2538 (Δ*comC*), expressing Strep-tag II-tagged BlpI (BlpI-strepII) were created in a similar fashion using an IDT-synthesized DNA fragment containing a Strep-tag II sequence (ggtggaggttcaggaggtggtagtgcttggtcacatccacaatttgaaaaa) in place of a HiBiT sequence.

Construction of P_{comAB}-luc *reporters.*

First, a strain containing the Janus cassette and *luc* inserted in the *bgaA* gene was created as follows. Four DNA fragments were PCR amplified: 1) from P654 using primers CW121/CW110, 2) from a Janus-containing strain using primers CW105/CW106, 3) from P1666 using primers CW186/CW161, 4) from a P654 derivative carrying a *lacZ* reporter integration in *bgaA* (replacing nucleotides 91-6582) followed by the *B. subtilis rrnB* and *S. pneumoniae rpsI* and *tufA* terminators (the sequences of the transcriptional terminators were obtained from ref (220)) using primers CW183/CW122. These fragments were ligated together using Gibson assembly and the crude product was PCR amplified using primers CW109/CW112. This final PCR product was used to transform P1613 to create strain P2166.

Next, a DNA fragment (1xterm-P_{comAB}-$luc_{5'}$) containing the *S. pneumoniae rpsD* terminator followed by the *comAB* promoter (see above) fused to the first 103 bp of *luc* was synthesized by IDT using their gBlocks service. Then, the following DNA fragments were PCR amplified: 1) from P654 using primers CW121/CW115, 2) from 1xterm-P_{comAB}-$luc_{5'}$ using primers CW117/CW190, 3) from P1666 using primers CW191/CW189. These fragments were ligated together using Gibson assembly and the crude product was PCR amplified using primers CW109/CW188. This final PCR product was used to transform P2166 to create strain P2187. P2187 was then transformed as described above to create the Δ*comC* mutant P2248.

Construction of tagged ComA.

An R6 strain expressing Twin-Strep-tagged ComA was created as follows. First, a DNA fragment containing an insertion of a codon-optimized Twin-Strep-tag (194) sequence preceded by a -GSSGGGGSGGGGS- linker (ggttcatcaggtggaggtggatcaggaggtggtggttcttcagcatggtcacaccctcaattcgaaaaaggtggaggatctggtggtggatcaggtggttcatctgcttggtcacatccacaatttgaaaaa) between the last coding triplet and the stop triplet of *comA* was synthesized by IDT using their gBlocks service. This fragment was PCR amplified using primers CW95/CW94. Next, upstream and downstream homology fragments were PCR amplified from P654 using primers CW268/CW309 and CW310/CW98. These three fragments were ligated together using Gibson assembly and the crude product was amplified using primers CW84/CW87. This final product was used to transform P2516 via Janus exchange with the *comAB*-Janus PCR product to create strain P2546.

Construction of mouse-adapted serotype 19A strains.

The spectinomycin-resistant serotype 19A “killer” strains encoding the P164-type BIR were created as follows. P213 (14) (referred to as 19*Blp*$_{164}$ in reference), a serotype 19A strain carrying an intact copy of *blpA* and BIR$_{P164}$, was transformed with a spectinomycin resistance cassette (*ant(9)-Ia*) insertion downstream of *blpT* as described in (40), then mouse passaged by nasopharyngeal colonization for 1-2 days using the same techniques as in the colonization assays described in Chapter II to create strain P1163. The BlpAB^{-} counterpart was created by introducing a frameshifted version of *blpA* to P213 via Janus exchange and transforming the subsequent strain with the spectinomycin resistance cassette and mouse passaging as above to create strain P2014. The Janus-containing PCR product used in the above transformation was amplified from PSD129 (183) using primers SD3/SD4. The frameshifted *blpA*-containing PCR product used in the above transformation was amplified from P174 (14) using primers SD3/SD4.

The spectinomycin-resistant serotype 19A strains encoding the P133-type BIR were created as follows. P1047 (14) (referred to as 19*Blp*$_{133}$C$_{6A}$ in reference) was made spectinomycin resistant and mouse passaged as above to create strain P2078, the BlpCH$_{6A}$-encoding BlpAB(+) strain. The BlpCH$_{R6}$-encoding counterpart was created by

making P694 (14) (referred to as 19*Blp*$_{133}$C$_{R6}$ in reference) spectinomycin resistant and mouse passaging the resulting strain as above to create strain P2200.

The kanamycin-resistant serotype 19A Δ*blp* pneumocin-sensitive strain was created by mouse passaging P690 (14) as above to create strain P824.

Construction of the Janus+ cassette.

Two PCR products, amplified using primer pairs CW105/CW357 on Janus cassette (188) template, and CW358/CW106 on Sweet Janus+ cassette template, were ligated via Gibson assembly. The resulting 1402-bp product formed the Janus+ cassette, which contains the *aphA-3* gene (kanamycin resistance) driven by the pneumococcal *amiA* promoter, followed by the *rpsL+* gene (dominant-negative streptomycin sensitivity) driven by the pneumococcal *rplM* promoter. The version of *rpsL+* included in Janus+ has the same five silent substitutions as the version in Sweet Janus+. These substitutions and the increased expression from the *rplM* promoter collectively reduce the rate of spontaneous reversion to a streptomycin resistant phenotype in strains carrying the Janus+ cassette compared to the original Janus cassette.

Construction of the Janus2 cassette.

A gBlocks DNA fragment was synthesized by IDT containing the gentamicin resistance gene *aacC1* from pPEPY-PF6-*lacI* (221). This template was PCR amplified using primers CW407/CW406 and ligated via Gibson assembly with another PCR product amplified using primers CW105/CW375 from Sweet Janus+ cassette template. The resulting 2069-bp product formed the Janus2 cassette, which contains the *sacB* and *aacC1* genes in a single operon driven by pneumococcal *amiA* promoter.

Construction of the Sp9-BS68 constitutive luciferase reporter.

The Sp9-BS68 PF6-*luc* reporter in which the firefly luciferase gene (*luc*) is expressed from the highly active PF6 promoter (221) was created as follows. First, a PCR product (Sp9-BS68-CEP-Janus2-*luc*) of the Janus2 cassette and *luc* inserted into the CEP site (222) between *treR* and *amiF* was amplified using primers CW303/CW295 following Gibson assembly of four PCR products amplified using primer pairs

CW456/CW454 on Sp9-BS68 template, CW105/CW406 on Janus2 cassette template, CW457/CW458 on P1666 (206) template, and CW343/CW294 on a GenParts fragment (Genscript) containing the *thrC* terminator from *E. coli* (223), the *tufA* terminator from pneumococcus (220), and approximately 500 bp of flanking sequence including *treR*. The assembled product contains a deletion of the region from n.t. +639 of the gene *CGSSp9BS68_00992* to n.t. −466 of the gene *CGSSp9BS68_00972*, into which are inserted the Janus2 cassette followed immediately by *luc*, oriented in the same direction as *amiF*, then a 50-n.t. spacer and finally the two terminators. The deletion removes an ABC transporter operon predicted to encode a non-functional product due to disruptions of the putative substrate-binding protein and permease genes. Second, a PCR product (Sp9-BS68-CEP-PF6-*luc*) of the luciferase reporter was amplified using primers CW463/CW188 following Gibson assembly of three PCR products amplified using primer pairs CW464/CW455 on Sp9-BS68 template, CW380/CW190 on a GenParts fragment (Genscript) containing the *hisI* and *rpsI* terminators from *E. coli* (223), the PF6 promoter, and the first 103 n.t. of *luc*, and CW191/CW189 on P1666 template. The Sp9-BS68-CEP-Janus2-*luc* PCR product was transformed into Sp9-BS68 to create strain P2769 and then the Janus2 cassette was exchanged with the Sp9-BS68-CEP-PF6-*luc* PCR product to create strain P2772.

Construction of Sp9-BS68 rtg luciferase reporters.

The Sp9-BS68 P_{rtgA}-*luc* reporter in which the firefly luciferase gene (*luc*) is inserted in place of *rtgA* following an ectopic copy of the *rtgS1*/*rtgA* promoter was created as follows. A PCR product (Sp9-BS68-CEP-P_{rtgA}-*luc*) was amplified using primers CW463/CW188 following Gibson assembly of three PCR products amplified using primer pairs CW464/CW455 on Sp9-BS68 template, CW380/CW190 on a GenParts fragment (Genscript) containing the *hisI* and *rpsI* terminators from *E. coli*, the region from n.t. −429 to n.t. −1 of *rtgA*, and the first 103 n.t. of *luc*, and CW191/CW189 on P1666 template. Then, the Janus2 cassette in P2769 was exchanged with the Sp9-BS68-CEP-P_{rtgA}-*luc* PCR product to create strain P2775. While the ectopic promoter region of *rtgA* in front of *luc* in P2775 contains a copy of *rtgS1*, this copy is disrupted by a frameshift mutation. Therefore, P2775 still only has one functional copy of *rtgS1*.

The Sp9-BS68 P_{rtgS1}-*luc* reporter in which the firefly luciferase gene (*luc*) is inserted in place of *rtgS1* following an ectopic copy of the *rtgS1* promoter was created as follows. A PCR product (Sp9-BS68-CEP-P_{rtgS1}-*luc*) was amplified using primers CW463/CW188 following Gibson assembly of two PCR products amplified using primer pairs CW464/CW494 and CW158/CW189 on P2775 template. Then, the Janus2 cassette in P2769 was exchanged with the Sp9-BS68-CEP-P_{rtgS1}-*luc* PCR product to create strain P2792.

Construction of D39 rtg luciferase reporter.

The D39 P_{rtgS1}-*luc* reporter in which the firefly luciferase gene (*luc*) is inserted in place of *rtgS1* following an ectopic copy of the *rtgS1* promoter was created as follows. First, a PCR product (D39-CEP-Janus2-*luc*) of the Janus2 cassette and *luc* inserted into the CEP site between *treR* and *amiF* was amplified using primers CW463/CW293 following Gibson assembly of three PCR products amplified using primer pairs CW464/CW471 on P2055 template, CW105/CW474 on P2769 template, and CW481/CW292 on P2055 template. The assembled product contains a deletion of the region from n.t. −101 to n.t. +1168 of the gene *SPD_1666*, replaced by the insertion from P2769. The deletion removes a degenerate transposon. Second, a PCR product (Sp9-BS68-CEP-P_{rtgS1}-*luc*) containing the luciferase reporter was amplified using primers CW463/CW188 following Gibson assembly of three PCR products amplified using primer pairs CW464/CW482 on P2055 template, CW380/CW494 on P2775 template, and CW158/CW189 on P1666 template. The D39-CEP-Janus2-*luc* PCR product was transformed into P2055 to create strain P2779 and then the Janus2 cassette was exchanged with the D39-CEP-P_{rtgS1}-*luc* PCR product to create strain P2790.

Construction of Sp9-BS68 rtgR and rtgS1 deletion mutants.

The Sp9-BS68 Δ*rtgR* strain was created as follows. First, a PCR product (Sp9-BS68-*rtgR*-Janus2) was amplified using primers CW487/CW486 following Gibson assembly of three PCR products amplified using primer pairs CW488/CW490 on Sp9-BS68 template, CW105/CW406 on Janus2 cassette template, and CW489/CW485 on Sp9-BS68 template. The assembled product contains the Janus2 cassette inserted in

place of *rtgR*. Second, a PCR product (Sp9-BS68-Δ*rtgR*) was amplified using primers CW487/CW486 following Gibson assembly of two PCR products amplified using primer pairs CW488/CW491 and CW492/CW485 on Sp9-BS68 template. The Sp9-BS68-*rtgR*-Janus2 PCR product was transformed into P2792 and then the Janus2 cassette was exchanged with the Sp9-BS68-Δ*rtgR* PCR product to create strain P2802.

The Sp9-BS68 Δ*rtgS1* strain was created as follows. First, a PCR product (Sp9-BS68-*rtgR*-*rtgS1*-Janus2) was amplified using primers CW498/CW486 following Gibson assembly of three PCR products amplified using primer pairs CW499/CW523 on Sp9-BS68 template, CW105/CW406 on Janus2 cassette template, and CW489/CW485 on Sp9-BS68 template. The assembled product contains the Janus2 cassette inserted in place of *rtgR* and *rtgS1*. Second, a PCR product (Sp9-BS68-Δ*rtgS1*) was amplified using primers CW498/CW486 following Gibson assembly of three PCR products amplified using primer pairs CW488/CW503 on Sp9-BS68 template, CW502/CW484 on a gBlocks fragment (IDT) containing an in-frame deletion of n.t. +31 to n.t. +84 of *rtgS1*, and CW483/CW485 on Sp9-BS68 template. The Sp9-BS68-*rtgR*-*rtgS1*-Janus2 PCR product was transformed into P2792 to create strain P2798 and then the Janus2 cassette was exchanged with the Sp9-BS68-Δ*rtgS1* PCR product to create strain P2804.

The Sp9-BS68 Δ*rtgR*Δ*rtgS1* strain was created as follows. A PCR product (Sp9-BS68-Δ*rtgR*Δ*rtgS1*) was amplified using primers CW498/CW486 following Gibson assembly of two PCR products amplified using primer pairs CW488/CW484 on P2804 template and CW483/CW485 on P2802 template. Then, the Janus2 cassette in P2798 was exchanged with the Sp9-BS68-Δ*rtgR*Δ*rtgS1* PCR product to create strain P2811.

Construction of D39 rtgS1 and rtgS2 deletion mutants.

The D39 Δ*rtgS1* strain was created as follows. First, a PCR product (D39-*rtgS1*-Janus2) was amplified using primers CW556/CW498 following Gibson assembly of three PCR products amplified using primer pairs CW496/CW500 on P2055 template, CW105/CW406 on Janus2 cassette template, and CW501/CW499 on P2055 template. Second, a PCR product (D39-Δ*rtgS1*) was amplified using primers CW556/CW498 following Gibson assembly of three PCR products amplified using primer pairs CW496/CW483 on P2055 template, CW484/CW502 on P2804 template, and

CW503/CW499 on P2055 template. The D39-*rtgS1*-Janus2 PCR product was transformed into P2790 to create strain P2851 and then the Janus2 cassette was exchanged with the D39-Δ*rtgS1* PCR product to create strain P2859.

The D39 Δ*rtgS2* strain was created as follows. First, a PCR product (D39-*rtgS2*-Janus+) was amplified using primers CW513/CW518 following Gibson assembly of three PCR products amplified using primer pairs CW548/CW562 on P2055 template, CW105/CW106 on Janus+ cassette template, and CW561/CW519 on P2055 template. Second, a PCR product (D39-Δ*rtgS2*) was amplified using primers CW513/CW518 following Gibson assembly of three PCR products amplified using primer pairs CW548/CW483 on P2055 template, CW484/CW520 on a gBlocks fragment (IDT) containing an in-frame deletion of n.t. +31 to n.t. +84 of *rtgS2*, and CW521/CW519 on P2055 template. Third, a PCR product (D39-*rtgS1*) was amplified using primers CW556/CW498 on P2055 template. The D39-*rtgS2*-Janus+ PCR product was transformed into P2851 to create strain P2896. Then, the Janus+ cassette was exchanged with the D39-Δ*rtgS2* PCR product to create strain P2902. Finally, the Janus2 cassette was exchanged with the D39-*rtgS1* PCR product to create strain P2908.

The D39 Δ*rtgS1*Δ*rtgS2* strain was created as follows. The Janus2 cassette in strain P2902 was exchanged with the D39-Δ*rtgS1* PCR product to create strain P2910.

Construction of Sp9-BS68 ami *and* ppt *deletion mutants.*

The Sp9-BS68 Δ*amiCD* strain was created as follows. First, a PCR product (Sp9-BS68-*amiCD*-Janus2) was amplified using primers CW616/CW607 following Gibson assembly of three PCR products amplified using primer pairs CW617/CW614 on Sp9-BS68 template, CW105/CW406 on Janus2 cassette template, and CW611/CW606 on Sp9-BS68 template. Second, a PCR product (Sp9-BS68-Δ*amiCD*) was amplified using primers CW616/CW607 following Gibson assembly of two PCR products amplified using primer pairs CW617/CW613 and CW612/CW606 on Sp9-BS68 template. The Sp9-BS68-*amiCD*-Janus2 PCR product was transformed into P2792, and then the Janus2 cassette was exchanged with the Sp9-BS68-Δ*amiCD* PCR product to create strain P3075.

The Sp9-BS68 Δ*pptAB* strain was created as follows. First, a PCR product (Sp9-BS68-*pptAB*-Janus2) was amplified using primers CW619/CW627 following Gibson assembly of three PCR products amplified using primer pairs CW618/CW622 on Sp9-BS68 template, CW105/CW406 on Janus2 cassette template, and CW625/CW628 on Sp9-BS68 template. Second, a PCR product (Sp9-BS68-Δ*amiCD*) was amplified using primers CW619/CW627 following Gibson assembly of two PCR products amplified using primer pairs CW618/CW623 and CW624/CW628 on Sp9-BS68 template. The Sp9-BS68-*pptAB*-Janus2 PCR product was transformed into P2792, and then the Janus2 cassette was exchanged with the Sp9-BS68-Δ*pptAB* PCR product to create strain P3077.

Construction of Sp9-BS68 rtgS1 promoter mutation reporters.

The Sp9-BS68 P_{rtgS1}(P2)-*luc* reporter strain was created as follows. First, the Sp9-BS68-CEP-Janus2-*luc* PCR product was amplified using primers CW463/CW468 on P2769 template. Second, a PCR product (Sp9-BS68-CEP-P_{rtgS1}(P2)-*luc*) was amplified using primers CW463/CW188 following Gibson assembly of two PCR products amplified using primer pairs CW464/CW570 and CW629/CW189 on P2792 template. The Sp9-BS68-CEP-Janus2-*luc* PCR product was transformed into strain P2804 to create strain P3080, and then the Janus2 cassette was exchanged with the Sp9-BS68-CEP-P_{rtgS1}(P2)-*luc* PCR product to create strain P3100.

The Sp9-BS68 P_{rtgS1}(P2)-*luc* reporter strains with promoter sequence mutations were created as follows. The mut1, mut2, mut3, and mut4 PCR products were amplified using primers CW463/CW188 following Gibson assembly of the following pairs of PCR products, all amplified from P3100 template: mut1, CW464/CW632 and CW633/CW189; mut2, CW464/CW630 and CW631/CW189; mut3, CW464/CW634 and CW635/CW189; mut4, CW464/CW636 and CW637/CW189. The Janus2 cassette in strain P3080 was then exchanged with mut1, mut2, mut3, and mut4 to create strains P3123, P3125, P3127, and P3129, respectively.

Construction of D39 com/blp luciferase reporters with repaired rtgA (RtgAB$^+$).

First, a PCR product (D39-*rtgAXB*-SJanus+) was amplified using primers CW556/CW557 following Gibson assembly of three PCR products amplified using

primer pairs CW496/CW537 on D39 template, CW105/CW106 on Sweet Janus+ cassette template, and CW509/CW559 on D39 template. Second, a PCR product (D39-*rtgAXB*$_{Sp9\text{-}BS68}$) was amplified using primers CW556/CW557 following Gibson assembly of three PCR products amplified using primer pairs CW496/CW502 on D39 template, CW503/CW511 on Sp9-BS68 template, and CW522/CW559 on D39 template. The D39-*rtgAXB*-SJanus+ PCR product was transformed into strains P2665, P2666, P2668, and P2670 and then the Sweet Janus+ cassettes were exchanged with the D39-*rtgAXB*$_{Sp9\text{-}BS68}$ PCR product to create strains P2838, P2840, P2842, and P2844, respectively.

Construction of Sp9-BS68 strains for mouse colonization assays.

The spectinomycin-resistant Sp9-BS68 strains were created as follows. A PCR product (Sp9-BS68-CEP-spcR) was amplified using primers CW463/CW468 following Gibson assembly of three PCR products amplified using primer pairs CW464/CW570 on P2792 template, CW571/CW572 on pE81 template, and CW343/CW292 on P2792 template. The Sp9-BS68-CEP-spcR PCR product was transformed into strains P2792 and P2811 to create strains P3001 and P3003, respectively.

The gentamicin-resistant Sp9-BS68 strain was created as follows. A PCR product (Sp9-BS68-CEP-genR) was amplified using primers CW463/CW468 following Gibson assembly of three PCR products amplified using primer pairs CW464/CW600 on P3001 template, CW598/CW599 on Janus2 cassette template, and CW343/CW292 on P3001 template. The Sp9-BS68-CEP-genR PCR product was transformed into strain P2811 to create strain P3025.

Mouse-passaged versions of the above strains were created as follows. P3001, P3003, and P3025 were inoculated into 5-7-week-old female BALB/c mice using the same protocol as that used for the single-strain colonization assays. After 24 hours, the colonizers were recovered via nasal wash using the same protocol as that used for the colonization assays. Nasal washes were plated on spectinomycin- or gentamicin-containing TSA plates supplemented with 5 µg/mL catalase and incubated overnight at 37°C with 5% CO_2. From each of the P3001, P3003, and P3025-inoculated mice, eight

individual colonies were pooled to create stocks of strains P3035, P3037, and P3039, respectively.

Construction of rtg autoinducing-deficient tagged BlpI expressing R6 strains.

First, a PCR product (R6-*rtgS1-rtgD2*-SJanus+) was amplified using primers CW556/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW496/CW500 on R6 template, CW105/CW106 on Sweet Janus+ cassette template, and CW538/CW551 on R6 template. Second, a PCR product (R6-Δ*rtgS1-rtgAXB*) was amplified using primers CW556/CW550 following Gibson assembly of two PCR products amplified using primers CW556/CW565 on P2859 template, and CW542/CW551 on R6 template. Third, a PCR product (R6-Δ*rtgS1-rtgAXB*$_{Sp9\text{-}BS68}$) was amplified using primers CW556/CW550 following Gibson assembly of three PCR products amplified using primers CW556/CW483 on R6 template, CW484/CW541 on P2804 template, and CW542/CW551 on R6 template. The R6-*rtgS1-rtgD2*-SJanus+ PCR product was transformed into strains P2538, P2565, P2567, and P2569 to create strains P2865, P2867, P2869, and P2871, respectively. Then, the Sweet Janus+ cassettes in P2865, P2867, P2869, and P2871 were exchanged for the R6-Δ*rtgS1-rtgAXB* PCR product to create strains P2878, P2880, P2884, and P2888, respectively. Finally, the Sweet Janus+ cassette in strain P2867 was exchanged for the R6-Δ*rtgS1-rtgAXB*$_{Sp9\text{-}BS68}$ PCR product to create strain P2882.

Construction of RtgC-HiBiT and RtgG-HiBiT expressing R6 strains.

The R6 strains expressing RtgC-HiBiT were created as follows. First, the pE57 insertion and *blpI*$_{P133}$*-HiBiT* genes were removed from P2880, P2882, P2884, and P2888. A PCR product (P654-*blpA-pncS*-Janus+) was amplified using primers CW271/CW219 following Gibson assembly of three PCR products amplified using primer pairs CW272/CW270 on P654 (183) template, CW105/CW106 on Janus+ cassette template, and CW238/CW234 on P654 template. Another PCR product (P654-ΔpE57) was amplified using primers CW585/CW219 following Gibson assembly of two PCR products amplified using primer pairs CW584/CW586 and CW587/CW234 on P654 template. The P654-*blpA-pncS*-Janus+ PCR product was transformed into strains P2880, P2882, P2884, and P2888 and then exchanged with the P654-ΔpE57 PCR

product to create strains P2934, P2936, P2938, and P2940, respectively. Second, *rtgC-HiBiT* was inserted downstream of *rtgB* in these strains. A PCR product (P2880-*rtgC-rtgD2*-SJanus+) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW531/CW534 on R6 template, CW105/CW106 on Sweet Janus+ cassette template, and CW538/CW551 on R6 template. Another PCR product (P2882-*rtgC-rtgD2*-SJanus+) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW543/CW546 on Sp9-BS68 template, CW105/CW106 on Sweet Janus+ cassette template, and CW538/CW551 on R6 template. A third PCR product (P2880-*rtgC-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW531/CW506 on R6 template, CW592/CW593 on a gBlocks fragment (IDT) containing $rtgC_{R6}$ with an insertion of sequence encoding HiBiT tag preceded by a 10-residue linker (ggtggtggaggttcaggaggtggaggttctgtttctggttggcgtctttttaaaaaaatttca) followed by $rtgD1_{R6}$, and CW594/CW551 on R6 template. A final PCR product (P2882-*rtgC-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW543/CW546 on a template containing *rtgAXB* from Sp9-BS68 inserted in place of *rtgAXB* in the D39 *rtg* locus, CW592/CW593 on the *rtgC-HiBiT* gBlocks fragment from above, and CW594/CW551 on R6 template. The P2880-*rtgC-rtgD2*-SJanus+ PCR product was transformed into strains P2934, P2938, and P2940 to create strains P2945, P2949, and P2951, respectively. The P2882-*rtgC-rtgD2*-SJanus+ PCR product was transformed into strain P2936 to create strain P2947. Then, the Sweet Janus+ cassettes in strains P2945, P2949, and P2951 were exchanged for the P2880-*rtgC-HiBiT* PCR product to create strains P2959, P2963, and P2965, respectively. Finally, the Sweet Janus+ cassette in strain P2947 was exchanged for the P2882-*rtgC-HiBiT* PCR product to create strain P2961.

The R6 strains expressing RtgG-HiBiT were created as follows. A PCR product (P2880-*rtgG-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW531/CW595 on R6 template, CW522/CW593 on a gBlocks fragment (IDT) containing $rtgG_{Sp9\text{-}BS68}$ with an insertion of sequence encoding HiBiT tag preceded by a 10-residue linker

(ggtggtggaggttcaggaggtggaggttctgtttctggttggcgtctttttaaaaaaatttca) followed by *rtgH*$_{Sp9\text{-}BS68}$, and CW594/CW551 on R6 template. Another PCR product (P2882-*rtgG-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW543/CW511 on P2882 template, CW522/CW593 on the *rtgG-HiBiT* gBlocks fragment from above, and CW594/CW551 on R6 template. The Sweet Janus+ cassettes in strains P2945, P2949, and P2951 were exchanged for the P2880-*rtgG-HiBiT* PCR product to create strains P2967, P2971, and P2973, respectively. The Sweet Janus+ cassette in strain P2947 was exchanged for the P2882-*rtgG-HiBiT* PCR product to create strain P2969.

Construction of the R6 strain with rtgA$_{ATG>ATT}$ *expressing RtgG-HiBiT.*

A PCR product (P2882-*rtgA*-Janus2) was amplified using primers CW556/CW659 following Gibson assembly of three PCR products amplified using primer pairs CW496/CW537 on P2882 template, CW105/CW406 on Janus2 cassette template, and CW657/CW660 on P2882 template. Another PCR product (P2882-*rtgA*$_{ATG>ATT}$) was amplified using primers CW556/CW659 following Gibson assembly of three PCR products amplified using primer pairs CW496/CW502 on P2882 template, CW503/CW446 on R6 template, and CW447/CW660 on P2882 template. The P2882-*rtgA*-Janus2 PCR product was transformed into P2969, and then the Janus2 cassette was exchanged with the P2882-*rtgA*$_{ATG>ATT}$ PCR product to create strain P3162.

Construction of R6 strains expressing BlpI-HiBiT from downstream of rtgB.

A PCR product (P2880-*blpI-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW531/CW595 on R6 template, CW522/CW596 on a gBlocks fragment (IDT) containing *blpI-HiBiT*, and CW597/CW551 on P2967 template. Another PCR product (P2882-*blpI-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW543/CW511 on P2882 template, CW522/CW596 on the *blpI-HiBiT* gBlocks fragment from above, and CW597/CW551 on P2967 template. The Sweet Janus+ cassettes in strains P2945, P2949, and P2951 were exchanged for the P2880-*blpI-HiBiT* PCR product to create

strains P3027, P3031, and P3033, respectively. The Sweet Janus+ cassette in strain P2947 was exchanged for the P2882-*blpI-HiBiT* PCR product to create strain P3029.

Construction of R6 strains expressing signal sequence-swapped RtgG-HiBiT and BlpI-HiBiT.

The R6 strains expressing SS_{RtgG}-BlpI-HiBiT were created as follows. A PCR product (P2880-SS_{rtgG}-*blpI-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW531/CW595 on R6 template, CW522/CW596 on a gBlocks fragment (IDT) containing SS_{rtgG}-*blpI-HiBiT*, and CW597/CW551 on P2967 template. Another PCR product (P2882-SS_{rtgG}-*blpI-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW543/CW511 on P2882 template, CW522/CW596 on the SS_{rtgG}-*blpI-HiBiT* gBlocks fragment from above, and CW597/CW551 on P2967 template. The Sweet Janus+ cassettes in strains P2945, P2949, and P2951 were exchanged for the P2880-SS_{rtgG}-*blpI-HiBiT* PCR product to create strains P2981, P2985, and P2987, respectively. The Sweet Janus+ cassette in strain P2947 was exchanged for the P2882-SS_{rtgG}-*blpI-HiBiT* PCR product to create strain P2983.

The R6 strains expressing SS_{BlpI}-RtgG-HiBiT were created as follows. A PCR product (P2880-SS_{blpI}-*rtgG-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW531/CW595 on R6 template, CW522/CW596 on a gBlocks fragment (IDT) containing SS_{blpI}-*rtgG-HiBiT*, and CW597/CW551 on P2967 template. Another PCR product (P2882-SS_{blpI}-*rtgG-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW543/CW511 on P2882 template, CW522/CW596 on the SS_{blpI}-*rtgG-HiBiT* gBlocks fragment from above, and CW597/CW551 on P2967 template. The Sweet Janus+ cassettes in strains P2945, P2949, and P2951 were exchanged for the P2880-SS_{blpI}-*rtgG-HiBiT* PCR product to create strains P2993, P2997, and P2999, respectively. The Sweet Janus+ cassette in strain P2947 was exchanged for the P2882-SS_{blpI}-*rtgG-HiBiT* PCR product to create strain P2995.

Construction of R6 strains expressing HiBiT-tagged BlpI cargo peptide fused to various signal sequences.

The R6 strains expressing SS_{RtgG}(F/M/L/V)-BlpI-HiBiT were created as follows. A PCR product (P2880-*SS_{rtgG}(F/M/L/V)-blpI-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW531/CW595 on R6 template, CW522/CW596 on a gBlocks fragment (IDT) containing *SS_{rtgG}(F/M/L/V)-blpI-HiBiT*, and CW597/CW551 on P2967 template. Another PCR product (P2882-*SS_{rtgG}(F/M/L/V)-blpI-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW543/CW511 on P2882 template, CW522/CW596 on the *SS_{rtgG}(F/M/L/V)-blpI-HiBiT* gBlocks fragment from above, and CW597/CW551 on P2967 template. The Sweet Janus+ cassettes in strains P2945, P2949, and P2951 were exchanged for the P2880-*SS_{rtgG}(F/M/L/V)-blpI-HiBiT* PCR product to create strains P3009, P3013, and P3015, respectively. The Sweet Janus+ cassette in strain P2947 was exchanged for the P2882-*SS_{rtgG}(F/M/L/V)-blpI-HiBiT* PCR product to create strain P3011.

The R6 strains expressing SS_{BlpI}(Y/L/M/L)-BlpI-HiBiT were created as follows. A PCR product (P2880-*SS_{blpI}(Y/L/M/L)-blpI-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW531/CW595 on R6 template, CW522/CW596 on a gBlocks fragment (IDT) containing *SS_{blpI}(Y/L/M/L)-blpI-HiBiT*, and CW597/CW551 on P2967 template. Another PCR product (P2882-*SS_{blpI}(Y/L/M/L)-blpI-HiBiT* was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW543/CW511 on P2882 template, CW522/CW596 on the *SS_{blpI}(Y/L/M/L)-blpI-HiBiT* gBlocks fragment from above, and CW597/CW551 on P2967 template. The Sweet Janus+ cassettes in strains P2945, P2949, and P2951 were exchanged for the P2880-*SS_{blpI}(Y/L/M/L)-blpI-HiBiT* PCR product to create strains P3017, P3021, and P3023, respectively. The Sweet Janus+ cassette in strain P2947 was exchanged for the P2882-*SS_{blpI}(Y/L/M/L)-blpI-HiBiT* PCR product to create strain P3019.

The R6 strains expressing $SS_{RtgG}(N6_{BlpI})$-BlpI-HiBiT were created as follows. A PCR product (P2880-*$SS_{rtgG}(N6_{blpI})$-blpI-HiBiT*) was amplified using primers

CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW531/CW595 on R6 template, CW522/CW596 on a gBlocks fragment (IDT) containing *$SS_{rtgG}(N6_{blpI})$-blpI-HiBiT*, and CW597/CW551 on P2967 template. Another PCR product (P2882-*$SS_{rtgG}(N6_{blpI})$-blpI-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW543/CW511 on P2882 template, CW522/CW596 on the *$SS_{rtgG}(N6_{blpI})$-blpI-HiBiT* gBlocks fragment from above, and CW597/CW551 on P2967 template. The Sweet Janus+ cassettes in strains P2945, P2949, and P2951 were exchanged for the P2880-*$SS_{rtgG}(N6_{blpI})$-blpI-HiBiT* PCR product to create strains P3055, P3059, and P3061, respectively. The Sweet Janus+ cassette in strain P2947 was exchanged for the P2882-*$SS_{rtgG}(N6_{blpI})$-blpI-HiBiT* PCR product to create strain P3057.

The R6 strains expressing $SS_{BlpI}(N6_{RtgG})$-BlpI-HiBiT were created as follows. A PCR product (P2880-*$SS_{blpI}(N6_{rtgG})$-blpI-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW531/CW595 on R6 template, CW522/CW596 on a gBlocks fragment (IDT) containing *$SS_{blpI}(N6_{rtgG})$-blpI-HiBiT*, and CW597/CW551 on P2967 template. Another PCR product (P2882-*$SS_{blpI}(N6_{rtgG})$-blpI-HiBiT*) was amplified using primers CW451/CW550 following Gibson assembly of three PCR products amplified using primer pairs CW543/CW511 on P2882 template, CW522/CW596 on the *$SS_{blpI}(N6_{rtgG})$-blpI-HiBiT* gBlocks fragment from above, and CW597/CW551 on P2967 template. The Sweet Janus+ cassettes in strains P2945, P2949, and P2951 were exchanged for the P2880-*$SS_{blpI}(N6_{rtgG})$-blpI-HiBiT* PCR product to create strains P3081, P3085, and P3087, respectively. The Sweet Janus+ cassette in strain P2947 was exchanged for the P2882-*$SS_{blpI}(N6_{rtgG})$-blpI-HiBiT* PCR product to create strain P3083.

APPENDIX B: Primers

Table B.1. Primer list.

Primer	*Sequence*
CW13	GATGGTATCGCAGAGTATTCAAAACG
CW14	atcaaacggaAAGCCGGGAAAATTCCCAGC
CW15	ttcccggcttTCCGTTTGATTTTTAATGGATAATGTG
CW16	tatttcattaCAGAGACCTGGGCCCCTTTC
CW19	ggaacagtttTAATGAAATAAGGGGAAAGAG
CW20	ttatttcattaAAACTGTTCCAATTTAACTGTGTTTTT
CW23	caggtctctgTAATGAAATAAGGGGAAAGAGTAATGG
CW24	CATTCCAGCATAATCATGTCG
CW33	GCACACGGGAAAAGTTGTCT
CW34	acgaacattaCATCTCCTTTTCCCTAATACTCAATG
CW35	aaaggagatgTAATGTTCGTGTTTTTAGAGTTAAATAATTTTTAAAC
CW36	GCAAACAAGTCCTCAGCCAA
CW39	TGAGCGCCAAACCATAGGAT
CW45	CAAACACAAGGAACGGCAGA
CW46	CGCTGTGTCGCTTGATCTAG
CW58	atcaaacggaTCAAGACGTTTCGATGCCAA
CW59	aacgtcttgaTCCGTTTGATTTTTAATGGATAATGTG
CW60	gacattctccCAGAGACCTGGGCCCCTTTC
CW61	caggtctctgGGAGAATGTCATGAATCCTAATCTTT
CW74	GCAAGCTGGTATCGGAAGTG
CW77	TGGAACGCTATAGTCGGCAT
CW78	ACTGTTCGCCATCTCATCCT
CW80	ATACCGTTATGAAAATTGTGGAAAG
CW81	TTTTGTGCCAGTAAGACGCC
CW82	CGGTAGTTCAGGCAGTTCAA
CW83	ACTACGCGTACTGTGAGCC
CW84	CGTAGTCCAGTTTGGCGATG
CW87	TTCAGCCTGCTCCGAAGTTA
CW92	TGAGATTGACACCACCCAGA
CW94	GGCCATGGGTACAATCACAC
CW95	TGATTTGCTTGCACAGGGAG
CW98	CGCCTTTCGTATTACTTGAGCA
CW102	ACAGCGGATGGTTCGGATAA
CW105	TCCGTTTGATTTTTAATGGATAATGTG
CW106	CAGAGACCTGGGCCCCTTTC
CW108	cccaggtctctgAAACAAGACCGAGAAACAAGAAC

Primer	*Sequence*
CW109	ACTGCCGGTGTATCTTTGGA
CW110	cattaaaaatcaaacggaTGCACAAGTCCCAATCATTACT
CW111	cccaggtctctgCCAAATACAGGTAGTGCGGC
CW112	ATCGCCCGTCTTATCACCTT
CW115	ccaaagtgcttctattaTGCACAAGTCCCAATCATTACT
CW117	TAATAGAAGCACTTTGGGACGTTCT
CW121	CGGTTTCATCCTTGGTATCGT
CW122	GAACATCACTTCTACGCCTTCT
CW131	cttGCTTAGTAGACCATTCTTTTTATCTTTATG
CW132	ggtctactaagcAAGCTTCCTTCCTCTGATTTTCA
CW143	CAATTTTCATAACGGTATAGAGACCAAC
CW144	CTTATCAAAAAGGAGAATCATAACATGG
CW145	TTCTCCTTTTTGATAAGATAATAAATAGTTATAGAG
CW147	TTAAAGTTGTAACAGTTGGGAAACT
CW148	ACCACTGTCAATTGCCATCTG
CW150	GCTCCACAGTTTCGGGTTTT
CW158	ATGGAAGACGCCAAAAACATAAAG
CW161	cggcgtcctactaTTACAATTTGGACTTTCCGCCC
CW165	CCCTACAAGTGTCATTGCATCT
CW167	CGATACCTTTACTTGGAGCGG
CW170	TTATTTGTCATCGTCGTCTTTGTAGTC
CW171	TCTCATAGCTCAGCTGGATAGAG
CW172	GCTGAGCTATGAGACCTAATACAATT
CW173	gacgatgacaaaTAATGAAATAAGGGGAAAGAGTAATGG
CW182	cattaaaaatcaaacggaATTCATCTCCTTTTCCCTAATACTCAAT
CW183	TAGTAGGACGCCGCCAAG
CW186	cccaggtctctgATGGAAGACGCCAAAAACATAAAG
CW188	GGAACAACTTTACCGACCGC
CW188	GGAACAACTTTACCGACCGC
CW189	CCCGGTATCCAGATCCACAA
CW190	GGGCGTATCTCTTCATAGCCT
CW191	GAAGAGATACGCCCTGGTTCC
CW194	cggcgtcctactaCTAGGCCAAAATACGTTCACACA
CW199	ggtggtgaactactATGGAAACTGAACGTGAAGAAAATG
CW200	aagttgttctgcttcCTAACCTCCTGCTTGTGGTTTC
CW201	AGTAGTTCACCACCTTTTCCCT
CW202	GAAGCAGAACAACTTTAACGCC
CW212	TCTCCGTCGCTATTGTAACCA
CW213	GAGTTTATCACCCTTGTCACTAAGA
CW219	TCAATCAGGACAGTCAAATCGA
CW232	cccaggtctctgTTGGCATCGAAACGTCTTGAA
CW234	TCTTCAAAAGTCGTGCGTTGA
CW238	cattaaaaatcaaacggaTCTTTTGCAAACGAGTCGCT
CW241	CAACGTCGTGAACATTCCGA

Primer	*Sequence*
CW242	TGGTGCATTTGTAGTCTTGCA
CW245	cattaaaaatcaaacggaACCTTGTAAGCCATATTGGATGAG
CW246	aagtgcTGAAGTTAGGTTTTGTTTCTTATCCATG
CW247	ctaacttcaGCACTTGAACTACCTATTCAGCTAT
CW255	TGTTTAGGATTGGGTGGGCT
CW256	TCCAACTAAAGCCCATACCGT
CW257	cccaggtctctgAAACTCTCCTAGCTCAGTGTAAGT
CW261	CAATACAACCACCGATTGCAGA
CW262	CGGTGGTTGTATTGGAGCAG
CW263	AACGTTCCAACCAGCCATTAG
CW264	TGGTTGGAACGTTAAAAATTCACTT
CW265	atccaaaaaatatacatAAACTCTCCTAGCTCAGTGTAAGT
CW266	TTTATGTATATTTTTTGGATTATATTGTATACACTTA
CW268	AGCACTTGAATCGACGCTTT
CW270	cccaggtctctgTGAACCAACAAACGACTTTTAGT
CW271	AGGTATGTAGGCGGTGCTAC
CW272	AACCCGGTAAGACACGACTT
CW292	TGGAATTGACTCGATAGCTTTAACA
CW293	AAAACAAACCGCATCCGTGT
CW294	TCCAACCTAACCAGCTACCA
CW295	CTGCGCTCTAAAACCAACGT
CW303	TCTGTACGTGCCCAAGTCTT
CW309	TGTGCAAGCAAATCAGCATG
CW310	TGTACCCATGGCCCTTCTG
CW343	TGAAAGCTGGCGCATGATG
CW357	GATATCCTCCCTGATCGACCG
CW358	ATCAGGGAGGATATCGGGGA
CW359	cccaggtctctgCTGGCTTCGACCATCCTTTC
CW360	TGTCTGAAGCTTGGTTTGCA
CW375	acttttcctccTTCTTTTGCGTTTTTATTTGTTAACTGT
CW376	agaaggaggaaaagtATGGCTAAAATGAGAATATCACCGG
CW380	TAATAGAACGCATGAGAAAGCCC
CW406	CGAGTTACTGGAGGGATCCTTA
CW407	agaaggaggaaaagtATGTTACGCAGCAGCAACG
CW446	AATACCACTGACAGTCGTTCCA
CW447	ACTGTCAGTGGTATTGTTTCAGG
CW451	AGAAGGTTTACAAGCGAGCTC
CW454	cattaaaaatcaaacggaCCACCTTTTTCATCACCTGTCAA
CW455	ctcatgcgttctattaCCACCTTTTTCATCACCTGTCAA
CW456	AGCACTTGACTCGTTACCCT
CW457	ctccagtaactcgATGGAAGACGCCAAAAACATAAAG
CW458	cgccagctttcaTTACAATTTGGACTTTCCGCCC
CW463	AGTAAGAAGTTTGTCGCGGTT
CW464	TGGCTGACTAGGAGGAAGGA

Primer	*Sequence*
CW468	ACGATACCAAACTTGTCTGCAA
CW471	cattaaaaatcaaacggaGGAAGCGACCAATAATCTCATCA
CW474	GTTAGACACTAAAAGAATCTTGCTTGG
CW481	ttcttttagtgtctaacCCCACTACAAATATTATAGAGCCGA
CW482	ctcatgcgttctattaGGAAGCGACCAATAATCTCATCA
CW483	ACCTACATAGTACCTCTTATTAAGCTCA
CW484	GAGGTACTATGTAGGTAGAAAATAAAGATAGG
CW485	TGAGGTTCAAAAAAGCGCCA
CW486	GACTTTCCCCATTGATCGCC
CW487	CACCTCGTCCATGATAACTTTTGA
CW488	CGAGATAGAAACAAAAGAACGTGC
CW489	ctccagtaactcgCTCTGCCAAGCCACATAGTG
CW490	cattaaaaatcaaacggaAGCAATGTAAAAAGTAAAAAAGAAAGTCGA
CW491	tagatctcTGACTCACGTATTTCTTTATAGACCTTT
CW492	tgagtcaGAGATCTATTACCAAGCGAAGGAAAG
CW494	tggcgtcttcCATACTTTTTCTCCTATCTTTATTTTCTACCTAC
CW496	TCCTCGGTTAGCTCCAGAAG
CW498	GGACAACCGCACTAATAACTGA
CW499	AGTCCATTTGCTTCCATTTGTTCA
CW500	cattaaaaatcaaacggaACCTACATAGTACCTCTTATTAAGCTCA
CW501	ctccagtaactcgGGCAGTGATTGAGGTTTGGG
CW502	CTCAATCACTGCCCAGTAGTTC
CW503	GGCAGTGATTGAGGTTTGGG
CW506	CAGCCCAATCAGCCCCATAT
CW509	cccaggtctctgGGAAATGTTAAGTTTCAAGCAATAATTGG
CW511	CTTGAAACTTAACATTTCTAAAATCCAATCAAG
CW513	CACTATGTGGCTTGGCAGAG
CW518	AGAATCTAATGACCGAGCTACTGT
CW519	ACACCTAAATAAAGTCATCTCTCCAA
CW520	GTGATCGCAGGCTTTGTTTAGT
CW521	GCCTGCGATCACTTGTTAGG
CW522	GAAATGTTAAGTTTCAAGCAATAATTGG
CW523	cattaaaaatcaaacggaGGCAGTGATTGAGGTTTGGG
CW531	CAGCAATTAGTTCAGGGAAGGT
CW534	cattaaaaatcaaacggaACAATCCGCACTTCAACTGG
CW537	cattaaaaatcaaacggaCCCAAACCTCAATCACTGCC
CW538	cccaggtctctgGCACAACATAGGGAGTTAGAGAAG
CW539	CTCCGCTTTTGTCTCTTCGT
CW541	agtccaaactattctTTATTGCTTGAAACTTAACATTTCTAAAATCC
CW542	AGAATAGTTTGGACTTGAACTGTATCTA
CW543	GCAATTGGTTCAGGGAAGGT
CW545	GGTACTGCTCGTTCAAACAACA
CW546	cattaaaaatcaaacggaTTATTGCTTGAAACTTAACATTTCTAAAATCC
CW548	TTGCTCGAATTTATAGGACTTTTTTCTAT

Primer	*Sequence*
CW550	GCAACAAAACCAGCCCCTAA
CW551	GCGGCTTTCTCCTCAGTTTT
CW556	AGACTTCTAGGCGTTGCAATATT
CW557	ATAGAAAAAAGTCCTATAAATTCGAGCAA
CW559	CGTTTATCCCAACTCAATTATGACATT
CW561	cattaaaaatcaaacggaTGAGTTGTTAGTGGAGGATAAAATGT
CW562	cccaggtctctgACCTACATAGTACCTCTTATTAAGCTCA
CW565	agtccaaactattctTTATTGCTTGAAACTTAACATTTCCAAAA
CW568	ACAAGTGATCCTCGACTGTGT
CW570	ACAATTTTCGAAAAAACCCGCTTC
CW571	gtttttcgaaaattgtCCCCCGTTTGATTTTTAATGGATAAT
CW572	cgccagctttcaCGCCCTTCAATTTTTTTATAATTTTTTTAATCT
CW584	GACCATTCTTTTTATCTTTATGGGGTT
CW585	TGGGTTTGGGAGCTAGAAAAATAG
CW586	ATCGAAACGTCTTGAATTAGCTTTTTTA
CW587	TCAAGACGTTTCGATGCCAA
CW592	GCTGATTGGGCTGACGGTA
CW593	TAGATACAGTTCAAGTCCAAACTATTCT
CW594	GACTTGAACTGTATCTAAAAACAGAGG
CW595	CTTGAAACTTAACATTTCCAAAATCCAATC
CW596	CCAACCAGAAACAGAACCTCC
CW597	CTGTTTCTGGTTGGCGTCTT
CW598	ATGTTACGCAGCAGCAACG
CW599	cgccagctttcaTTAGGTGGCGGTACTTGGG
CW600	tgctgcgtaacatATATATCCTCCTCACTATTTTGATTAGTACC
CW606	TTTCCCACTACCGGACTCAC
CW607	CCTTCCTCCTAGTCAGCCAG
CW611	cattaaaaatcaaacggaCGGCGTTTCCCTTGAACTAG
CW612	taccttgTCCCTTTTCGTAGTTGGTCAAAA
CW613	aaaagggaCAAGGTATAGATAATCGTGTAGGTCAA
CW614	ctccagtaactcgCCCAAAAGAACCTTCTCCTCATG
CW616	GCCTTGAATTTTGCTCTTGATCG
CW617	TGCTGGTCTGGGTGTGAATA
CW618	GAGTTACCTAAATTATGATGCATAGTTGA
CW619	GAGGATTTTAAAGTAATCTCTAACAATGCT
CW622	cattaaaaatcaaacggaAGGAACATGAACATAGCCACCT
CW623	aaccagcAGGAACATGAACATAGCCACCT
CW624	gttcctGCTGGTTTGGTTTTACTAGTCTTG
CW625	ctccagtaactcgCATTGCTGATACGACACTAAAAAAGA
CW627	GGCATACTTGACAGCGTATTGA
CW628	AGTTTCTCAATCGTCGTTTCAATTC
CW629	gtttttcgaaaattgtGCAATGTAAAAAGTAAAAAAGAAAGTCGA
CW630	ctttttttccaaaATTTATTTCGACTTTCTTTTTTACTTTTTACATTG
CW631	taaattttgGAAAAAAGTTAACGTAATTTTAAAATGAGCTT

Primer	*Sequence*
CW632	cttttttcacccATTTATTTCGACTTTCTTTTTTACTTTTTACATTG
CW633	taaatgggtGAAAAAAGTTAACGTAATTTTAAAATGAGCTT
CW634	aaaattcagtTAACTTTTTTCACAAATTTATTTCGACTTTCT
CW635	aaagttaactgAATTTTAAAATGAGCTTAATAAGAGGTACTATGT
CW636	cctcttaggAAGCTCATTTTAAAATTACGTTAACTTTTTTC
CW637	agcttcctaaGAGGTACTATGTAGGTAGAAAATAAAGATAGG
CW640	GTGGGGTATTAGGGGAGCAA
CW657	ctccagtaactcgTCAAAAGTTATCATGGACGAGGTG
CW659	CGAATTTTCCCCTCACTCGG
CW660	TGTCCCTGTAAAGAGTTCTATGGT
CW665	TGATACAGGAATGCAACAATAACTGA
CW666	TGCTTGAAGTGTATATAGGGATCAAA
CW668	CAATAACAACTGGCAAGAAAAGGT
SD1	GCGGGATCCCCAGCTTGAATTGATACACTAATGCTTTTATATAGGGAAAAGGTGGTGAACTACTATGG AAGACGCCAAAAACAT
SD2	GCGAGAGGCCTTTACAATTTGGAC
SD3	CTTATAAACGTACATTTGTTCC
SD4	GGTAACGTCTCTGATAAAAC

REFERENCES

1. Economou A. 2002. Bacterial secretome: the assembly manual and operating instructions (Review). Mol Membr Biol 19:159-69.
2. Green ER, Mecsas J. 2016. Bacterial Secretion Systems: An Overview. Microbiol Spectr 4.
3. Rees DC, Johnson E, Lewinson O. 2009. ABC transporters: the power to change. Nat Rev Mol Cell Biol 10:218-27.
4. Håvarstein LS, Diep DB, Nes IF. 1995. A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export. Mol Microbiol 16:229-40.
5. Ishii S, Yano T, Ebihara A, Okamoto A, Manzoku M, Hayashi H. 2010. Crystal structure of the peptidase domain of Streptococcus ComA, a bifunctional ATP-binding cassette transporter involved in the quorum-sensing pathway. J Biol Chem 285:10777-85.
6. Lin DY, Huang S, Chen J. 2015. Crystal structures of a polypeptide processing and secretion transporter. Nature 523:425-30.
7. Bobeica SC, Dong SH, Huo L, Mazo N, McLaughlin MI, Jiménez-Osés G, Nair SK, van der Donk WA. 2019. Insights into AMS/PCAT transporters from biochemical and structural characterization of a double Glycine motif protease. Elife 8.
8. Kotake Y, Ishii S, Yano T, Katsuoka Y, Hayashi H. 2008. Substrate recognition mechanism of the peptidase domain of the quorum-sensing-signal-producing ABC transporter ComA from Streptococcus. Biochemistry 47:2531-8.
9. Kanonenberg K, Schwarz CK, Schmitt L. 2013. Type I secretion systems - a story of appendices. Res Microbiol 164:596-604.
10. Gebhard S. 2012. ABC transporters of antimicrobial peptides in Firmicutes bacteria - phylogeny, function and regulation. Mol Microbiol 86:1295-317.
11. Chikindas ML, Weeks R, Drider D, Chistyakov VA, Dicks LM. 2018. Functions and emerging applications of bacteriocins. Curr Opin Biotechnol 49:23-28.
12. Riley MA, Wertz JE. 2002. Bacteriocins: evolution, ecology, and application. Annu Rev Microbiol 56:117-37.
13. Chassaing B, Cascales E. 2018. Antibacterial Weapons: Targeted Destruction in the Microbiota. Trends Microbiol 26:329-338.
14. Son MR, Shchepetov M, Adrian PV, Madhi SA, de Gouveia L, von Gottberg A, Klugman KP, Weiser JN, Dawid S. 2011. Conserved mutations in the pneumococcal bacteriocin transporter gene, blpA, result in a complex population consisting of producers and cheaters. MBio 2.
15. Mendonca ML, Szamosi JC, Lacroix AM, Fontes ME, Bowdish DM, Surette MG. 2016. The *sil* Locus in *Streptococcus* Anginosus Group: Interspecies Competition and a Hotspot of Genetic Diversity. Front Microbiol 7:2156.

16. Hale JD, Heng NC, Jack RW, Tagg JR. 2005. Identification of nlmTE, the locus encoding the ABC transport system required for export of nonlantibiotic mutacins in Streptococcus mutans. J Bacteriol 187:5036-9.
17. O'Keeffe T, Hill C, Ross RP. 1999. Characterization and heterologous expression of the genes encoding enterocin a production, immunity, and regulation in Enterococcus faecium DPC1146. Appl Environ Microbiol 65:1506-15.
18. Ishibashi N, Himeno K, Masuda Y, Perez RH, Iwatani S, Zendo T, Wilaipun P, Leelawatcharamas V, Nakayama J, Sonomoto K. 2014. Gene cluster responsible for secretion of and immunity to multiple bacteriocins, the NKR-5-3 enterocins. Appl Environ Microbiol 80:6647-55.
19. Birri DJ, Brede DA, Forberg T, Holo H, Nes IF. 2010. Molecular and genetic characterization of a novel bacteriocin locus in Enterococcus avium isolates from infants. Appl Environ Microbiol 76:483-92.
20. Axelsson L, Holck A. 1995. The genes involved in production of and immunity to sakacin A, a bacteriocin from Lactobacillus sake Lb706. J Bacteriol 177:2125-37.
21. Furgerson Ihnken LA, Chatterjee C, van der Donk WA. 2008. In vitro reconstitution and substrate specificity of a lantibiotic protease. Biochemistry 47:7352-63.
22. Ishibashi N, Zendo T, Koga S, Shigeri Y, Sonomoto K. 2015. Molecular characterization of the genes involved in the secretion and immunity of lactococcin Q, a two-peptide bacteriocin produced by Lactococcus lactis QU 4. Microbiology 161:2069-78.
23. Nishie M, Shioya K, Nagao J, Jikuya H, Sonomoto K. 2009. ATP-dependent leader peptide cleavage by NukT, a bifunctional ABC transporter, during lantibiotic biosynthesis. J Biosci Bioeng 108:460-4.
24. de Saizieu A, Gardès C, Flint N, Wagner C, Kamber M, Mitchell TJ, Keck W, Amrein KE, Lange R. 2000. Microarray-based identification of a novel Streptococcus pneumoniae regulon controlled by an autoinduced peptide. J Bacteriol 182:4696-703.
25. Perez RH, Himeno K, Ishibashi N, Masuda Y, Zendo T, Fujita K, Wilaipun P, Leelawatcharamas V, Nakayama J, Sonomoto K. 2012. Monitoring of the multiple bacteriocin production by Enterococcus faecium NKR-5-3 through a developed liquid chromatography and mass spectrometry-based quantification system. J Biosci Bioeng 114:490-6.
26. Johnston C, Martin B, Fichant G, Polard P, Claverys JP. 2014. Bacterial transformation: distribution, shared mechanisms and divergent control. Nat Rev Microbiol 12:181-96.
27. Hui FM, Morrison DA. 1991. Genetic transformation in Streptococcus pneumoniae: nucleotide sequence analysis shows comA, a gene required for competence induction, to be a member of the bacterial ATP-dependent transport protein family. J Bacteriol 173:372-81.
28. Håvarstein LS, Coomaraswamy G, Morrison DA. 1995. An unmodified heptadecapeptide pheromone induces competence for genetic transformation in Streptococcus pneumoniae. Proc Natl Acad Sci U S A 92:11140-4.
29. Martin B, Quentin Y, Fichant G, Claverys JP. 2006. Independent evolution of competence regulatory cascades in streptococci? Trends Microbiol 14:339-45.

30. Petersen FC, Scheie AA. 2000. Genetic transformation in Streptococcus mutans requires a peptide secretion-like apparatus. Oral Microbiol Immunol 15:329-34.
31. Reck M, Tomasch J, Wagner-Döbler I. 2015. The Alternative Sigma Factor SigX Controls Bacteriocin Synthesis and Competence, the Two Quorum Sensing Regulated Traits in Streptococcus mutans. PLoS Genet 11:e1005353.
32. Cuevas RA, Eutsey R, Kadam A, West-Roberts JA, Woolford CA, Mitchell AP, Mason KM, Hiller NL. 2017. A novel streptococcal cell-cell communication peptide promotes pneumococcal virulence and biofilm formation. Mol Microbiol 105:554-571.
33. Aggarwal SD, Eutsey R, West-Roberts J, Domenech A, Xu W, Abdullah IT, Mitchell AP, Veening JW, Yesilkaya H, Hiller NL. 2018. Function of BriC peptide in the pneumococcal competence and virulence portfolio. PLoS Pathog 14:e1007328.
34. Håvarstein LS, Holo H, Nes IF. 1994. The leader peptide of colicin V shares consensus sequences with leader peptides that are common among peptide bacteriocins produced by gram-positive bacteria. Microbiology 140 (Pt 9):2383-9.
35. Dobson AE, Sanozky-Dawes RB, Klaenhammer TR. 2007. Identification of an operon and inducing peptide involved in the production of lactacin B by Lactobacillus acidophilus. J Appl Microbiol 103:1766-78.
36. Nagao J, Morinaga Y, Islam MR, Asaduzzaman SM, Aso Y, Nakayama J, Sonomoto K. 2009. Mapping and identification of the region and secondary structure required for the maturation of the nukacin ISK-1 prepeptide. Peptides 30:1412-20.
37. Nishie M, Sasaki M, Nagao J, Zendo T, Nakayama J, Sonomoto K. 2011. Lantibiotic transporter requires cooperative functioning of the peptidase domain and the ATP binding domain. J Biol Chem 286:11163-9.
38. Sushida H, Ishibashi N, Zendo T, Wilaipun P, Leelawatcharamas V, Nakayama J, Sonomoto K. 2018. Evaluation of leader peptides that affect the secretory ability of a multiple bacteriocin transporter, EnkT. J Biosci Bioeng.
39. Kjos M, Miller E, Slager J, Lake FB, Gericke O, Roberts IS, Rozen DE, Veening JW. 2016. Expression of Streptococcus pneumoniae Bacteriocins Is Induced by Antibiotics via Regulatory Interplay with the Competence System. PLoS Pathog 12:e1005422.
40. Wholey WY, Kochan TJ, Storck DN, Dawid S. 2016. Coordinated Bacteriocin Expression and Competence in Streptococcus pneumoniae Contributes to Genetic Adaptation through Neighbor Predation. PLoS Pathog 12:e1005413.
41. Zendo T, Koga S, Shigeri Y, Nakayama J, Sonomoto K. 2006. Lactococcin Q, a novel two-peptide bacteriocin produced by Lactococcus lactis QU 4. Appl Environ Microbiol 72:3383-9.
42. Brömme D. 2001. Papain-like cysteine proteases. Curr Protoc Protein Sci Chapter 21:Unit 21.2.
43. Oldham ML, Chen J. 2011. Crystal structure of the maltose transporter in a pretranslocation intermediate state. Science 332:1202-5.
44. Zheng S, Nagao JI, Nishie M, Zendo T, Sonomoto K. 2018. ATPase activity regulation by leader peptide processing of ABC transporter maturation and

secretion protein, NukT, for lantibiotic nukacin ISK-1. Appl Microbiol Biotechnol 102:763-772.
45. Aniansson G, Alm B, Andersson B, Larsson P, Nylén O, Peterson H, Rignér P, Svanborg M, Svanborg C. 1992. Nasopharyngeal colonization during the first year of life. J Infect Dis 165 Suppl 1:S38-42.
46. Könönen E, Jousimies-Somer H, Bryk A, Kilp T, Kilian M. 2002. Establishment of streptococci in the upper respiratory tract: longitudinal changes in the mouth and nasopharynx up to 2 years of age. J Med Microbiol 51:723-30.
47. Turner P, Hinds J, Turner C, Jankhot A, Gould K, Bentley SD, Nosten F, Goldblatt D. 2011. Improved detection of nasopharyngeal cocolonization by multiple pneumococcal serotypes by use of latex agglutination or molecular serotyping by microarray. J Clin Microbiol 49:1784-9.
48. Said MA, Johnson HL, Nonyane BA, Deloria-Knoll M, O'Brien KL, Andreo F, Beovic B, Blanco S, Boersma WG, Boulware DR, Butler JC, Carratalà J, Chang FY, Charles PG, Diaz AA, Domínguez J, Ehara N, Endeman H, Falcó V, Falguera M, Fukushima K, Garcia-Vidal C, Genne D, Guchev IA, Gutierrez F, Hernes SS, Hoepelman AI, Hohenthal U, Johansson N, Kolek V, Kozlov RS, Lauderdale TL, Mareković I, Masiá M, Matta MA, Miró Ò, Murdoch DR, Nuermberger E, Paolini R, Perelló R, Snijders D, Plečko V, Sordé R, Strålin K, van der Eerden MM, Vila-Corcoles A, Watt JP, Team AAPBS. 2013. Estimating the burden of pneumococcal pneumonia among adults: a systematic review and meta-analysis of diagnostic techniques. PLoS One 8:e60273.
49. O'Brien KL, Wolfson LJ, Watt JP, Henkle E, Deloria-Knoll M, McCall N, Lee E, Mulholland K, Levine OS, Cherian T, Team HaPGBoDS. 2009. Burden of disease caused by Streptococcus pneumoniae in children younger than 5 years: global estimates. Lancet 374:893-902.
50. Cillóniz C, Amaro R, Torres A. 2016. Pneumococcal vaccination. Curr Opin Infect Dis 29:187-96.
51. Lucero MG, Dulalia VE, Nillos LT, Williams G, Parreño RA, Nohynek H, Riley ID, Makela H. 2009. Pneumococcal conjugate vaccines for preventing vaccine-type invasive pneumococcal disease and X-ray defined pneumonia in children less than two years of age. Cochrane Database Syst Rev:CD004977.
52. Moore MR, Link-Gelles R, Schaffner W, Lynfield R, Lexau C, Bennett NM, Petit S, Zansky SM, Harrison LH, Reingold A, Miller L, Scherzinger K, Thomas A, Farley MM, Zell ER, Taylor TH, Pondo T, Rodgers L, McGee L, Beall B, Jorgensen JH, Whitney CG. 2015. Effect of use of 13-valent pneumococcal conjugate vaccine in children on invasive pneumococcal disease in children and adults in the USA: analysis of multisite, population-based surveillance. Lancet Infect Dis 15:301-9.
53. Camilli R, D'Ambrosio F, Del Grosso M, Pimentel de Araujo F, Caporali MG, Del Manso M, Gherardi G, D'Ancona F, Pantosti A, Group PS. 2017. Impact of pneumococcal conjugate vaccine (PCV7 and PCV13) on pneumococcal invasive diseases in Italian children and insight into evolution of pneumococcal population structure. Vaccine 35:4587-4593.
54. Gaviria-Agudelo CL, Jordan-Villegas A, Garcia C, McCracken GH. 2017. The Effect of 13-Valent Pneumococcal Conjugate Vaccine on the Serotype

Distribution and Antibiotic Resistance Profiles in Children With Invasive Pneumococcal Disease. J Pediatric Infect Dis Soc 6:253-259.
55. Zuccotti G, Mameli C, Daprai L, Garlaschi ML, Dilillo D, Bedogni G, Faccini M, Gramegna M, Torresani E, Ballerini E, Benincaso A, Bonvissuto M, Bricalli D, Brioschi M, Calloni CS, Camiletti MI, Colella G, De Angelis L, Decarlis S, Di Nello F, Dozzi M, Galli E, Gandini V, Giuliani MG, Laviola F, Loda B, Macedoni M, Mazzucchi E, Metta MG, Moscatiello A, Nannini P, Petruzzi M, Picicco D, Picciotti M, Pisanelli S, Porta N, Ramponi G, Redaelli F, Rubini R, Sala N, Saitta V, Scelza G, Tiso RM, Tomasetto M, Torcoletti M, Travaini M, Valentini M, Vessia C, (PMSG) PSG. 2014. Serotype distribution and antimicrobial susceptibilities of nasopharyngeal isolates of Streptococcus pneumoniae from healthy children in the 13-valent pneumococcal conjugate vaccine era. Vaccine 32:527-34.
56. Lee GM, Kleinman K, Pelton SI, Hanage W, Huang SS, Lakoma M, Dutta-Linn M, Croucher NJ, Stevenson A, Finkelstein JA. 2014. Impact of 13-Valent Pneumococcal Conjugate Vaccination on Streptococcus pneumoniae Carriage in Young Children in Massachusetts. J Pediatric Infect Dis Soc 3:23-32.
57. Valente C, Hinds J, Gould KA, Pinto FR, de Lencastre H, Sá-Leão R. 2016. Impact of the 13-valent pneumococcal conjugate vaccine on Streptococcus pneumoniae multiple serotype carriage. Vaccine 34:4072-8.
58. Lee GM, Kleinman K, Pelton S, Lipsitch M, Huang SS, Lakoma M, Dutta-Linn M, Rett M, Hanage WP, Finkelstein JA. 2017. Immunization, Antibiotic Use, and Pneumococcal Colonization Over a 15-Year Period. Pediatrics 140.
59. Daana M, Rahav G, Hamdan A, Thalji A, Jaar F, Abdeen Z, Jaber H, Goral A, Huppert A, Raz M, Regev-Yochay G, group Ps. 2015. Measuring the effects of pneumococcal conjugate vaccine (PCV7) on Streptococcus pneumoniae carriage and antibiotic resistance: the Palestinian-Israeli Collaborative Research (PICR). Vaccine 33:1021-6.
60. Kim L, McGee L, Tomczyk S, Beall B. 2016. Biological and Epidemiological Features of Antibiotic-Resistant Streptococcus pneumoniae in Pre- and Post-Conjugate Vaccine Eras: a United States Perspective. Clin Microbiol Rev 29:525-52.
61. Cherazard R, Epstein M, Doan TL, Salim T, Bharti S, Smith MA. 2017. Antimicrobial Resistant Streptococcus pneumoniae: Prevalence, Mechanisms, and Clinical Implications. Am J Ther 24:e361-e369.
62. Hauser C, Kronenberg A, Allemann A, Mühlemann K, Hilty M. 2016. Serotype/serogroup-specific antibiotic non-susceptibility of invasive and non-invasive Streptococcus pneumoniae, Switzerland, 2004 to 2014. Euro Surveill 21.
63. Hanke CR, Grijalva CG, Chochua S, Pletz MW, Hornberg C, Edwards KM, Griffin MR, Verastegui H, Gil AI, Lanata CF, Klugman KP, Vidal JE. 2016. Bacterial Density, Serotype Distribution and Antibiotic Resistance of Pneumococcal Strains from the Nasopharynx of Peruvian Children Before and After Pneumococcal Conjugate Vaccine 7. Pediatr Infect Dis J 35:432-9.

64. Chen C, Cervero Liceras F, Flasche S, Sidharta S, Yoong J, Sundaram N, Jit M. 2019. Effect and cost-effectiveness of pneumococcal conjugate vaccination: a global modelling analysis. Lancet Glob Health 7:e58-e67.
65. Geno KA, Gilbert GL, Song JY, Skovsted IC, Klugman KP, Jones C, Konradsen HB, Nahm MH. 2015. Pneumococcal Capsules and Their Types: Past, Present, and Future. Clin Microbiol Rev 28:871-99.
66. Pai R, Moore MR, Pilishvili T, Gertz RE, Whitney CG, Beall B, Team ABCS. 2005. Postvaccine genetic structure of Streptococcus pneumoniae serotype 19A from children in the United States. J Infect Dis 192:1988-95.
67. Olarte L, Kaplan SL, Barson WJ, Romero JR, Lin PL, Tan TQ, Hoffman JA, Bradley JS, Givner LB, Mason EO, Hultén KG. 2017. Emergence of Multidrug-Resistant Pneumococcal Serotype 35B among Children in the United States. J Clin Microbiol 55:724-734.
68. Griffith F. 1928. The Significance of Pneumococcal Types. J Hyg (Lond) 27:113-59.
69. Avery OT, Macleod CM, McCarty M. 1944. STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES : INDUCTION OF TRANSFORMATION BY A DESOXYRIBONUCLEIC ACID FRACTION ISOLATED FROM PNEUMOCOCCUS TYPE III. J Exp Med 79:137-58.
70. Laurenceau R, Péhau-Arnaudet G, Baconnais S, Gault J, Malosse C, Dujeancourt A, Campo N, Chamot-Rooke J, Le Cam E, Claverys JP, Fronzes R. 2013. A type IV pilus mediates DNA binding during natural transformation in Streptococcus pneumoniae. PLoS Pathog 9:e1003473.
71. Méjean V, Claverys JP. 1993. DNA processing during entry in transformation of Streptococcus pneumoniae. J Biol Chem 268:5594-9.
72. Attaiech L, Olivier A, Mortier-Barrière I, Soulet AL, Granadel C, Martin B, Polard P, Claverys JP. 2011. Role of the single-stranded DNA-binding protein SsbB in pneumococcal transformation: maintenance of a reservoir for genetic plasticity. PLoS Genet 7:e1002156.
73. Mortier-Barrière I, Velten M, Dupaigne P, Mirouze N, Piétrement O, McGovern S, Fichant G, Martin B, Noirot P, Le Cam E, Polard P, Claverys JP. 2007. A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. Cell 130:824-36.
74. Mortier-Barrière I, de Saizieu A, Claverys JP, Martin B. 1998. Competence-specific induction of recA is required for full recombination proficiency during transformation in Streptococcus pneumoniae. Mol Microbiol 27:159-70.
75. Majewski J, Zawadzki P, Pickerill P, Cohan FM, Dowson CG. 2000. Barriers to genetic exchange between bacterial species: Streptococcus pneumoniae transformation. J Bacteriol 182:1016-23.
76. Sisco KL, Smith HO. 1979. Sequence-specific DNA uptake in Haemophilus transformation. Proc Natl Acad Sci U S A 76:972-6.
77. Goodman SD, Scocca JJ. 1988. Identification and arrangement of the DNA sequence recognized in specific transformation of Neisseria gonorrhoeae. Proc Natl Acad Sci U S A 85:6982-6.

78. Johnston C, Polard P, Claverys JP. 2013. The DpnI/DpnII pneumococcal system, defense against foreign attack without compromising genetic exchange. Mob Genet Elements 3:e25582.
79. Croucher NJ, Harris SR, Fraser C, Quail MA, Burton J, van der Linden M, McGee L, von Gottberg A, Song JH, Ko KS, Pichon B, Baker S, Parry CM, Lambertsen LM, Shahinas D, Pillai DR, Mitchell TJ, Dougan G, Tomasz A, Klugman KP, Parkhill J, Hanage WP, Bentley SD. 2011. Rapid pneumococcal evolution in response to clinical interventions. Science 331:430-4.
80. Croucher NJ, Chewapreecha C, Hanage WP, Harris SR, McGee L, van der Linden M, Song JH, Ko KS, de Lencastre H, Turner C, Yang F, Sá-Leão R, Beall B, Klugman KP, Parkhill J, Turner P, Bentley SD. 2014. Evidence for soft selective sweeps in the evolution of pneumococcal multidrug resistance and vaccine escape. Genome Biol Evol 6:1589-602.
81. Golubchik T, Brueggemann AB, Street T, Gertz RE, Spencer CC, Ho T, Giannoulatou E, Link-Gelles R, Harding RM, Beall B, Peto TE, Moore MR, Donnelly P, Crook DW, Bowden R. 2012. Pneumococcal genome sequencing tracks a vaccine escape variant formed through a multi-fragment recombination event. Nat Genet 44:352-5.
82. Croucher NJ, Harris SR, Barquist L, Parkhill J, Bentley SD. 2012. A high-resolution view of genome-wide pneumococcal transformation. PLoS Pathog 8:e1002745.
83. Cowley LA, Petersen FC, Junges R, Jimson D Jimenez M, Morrison DA, Hanage WP. 2018. Evolution via recombination: Cell-to-cell contact facilitates larger recombination events in Streptococcus pneumoniae. PLoS Genet 14:e1007410.
84. Zaccaria E, Wells JM, van Baarlen P. 2016. Metabolic Context of the Competence-Induced Checkpoint for Cell Replication in Streptococcus suis. PLoS One 11:e0153571.
85. Bergé MJ, Mercy C, Mortier-Barrière I, VanNieuwenhze MS, Brun YV, Grangeasse C, Polard P, Campo N. 2017. A programmed cell division delay preserves genome integrity during natural genetic transformation in Streptococcus pneumoniae. Nat Commun 8:1621.
86. Hui FM, Zhou L, Morrison DA. 1995. Competence for genetic transformation in Streptococcus pneumoniae: organization of a regulatory locus with homology to two lactococcin A secretion genes. Gene 153:25-31.
87. Håvarstein LS, Gaustad P, Nes IF, Morrison DA. 1996. Identification of the streptococcal competence-pheromone receptor. Mol Microbiol 21:863-9.
88. Alloing G, Martin B, Granadel C, Claverys JP. 1998. Development of competence in Streptococcus pneumonaie: pheromone autoinduction and control of quorum sensing by the oligopeptide permease. Mol Microbiol 29:75-83.
89. Martin B, Soulet AL, Mirouze N, Prudhomme M, Mortier-Barrière I, Granadel C, Noirot-Gros MF, Noirot P, Polard P, Claverys JP. 2013. ComE/ComE~P interplay dictates activation or extinction status of pneumococcal X-state (competence). Mol Microbiol 87:394-411.
90. Ween O, Gaustad P, Håvarstein LS. 1999. Identification of DNA binding sites for ComE, a key regulator of natural competence in Streptococcus pneumoniae. Mol Microbiol 33:817-27.

91. Boudes M, Sanchez D, Graille M, van Tilbeurgh H, Durand D, Quevillon-Cheruel S. 2014. Structural insights into the dimerization of the response regulator ComE from Streptococcus pneumoniae. Nucleic Acids Res 42:5302-13.
92. Martin B, Prudhomme M, Alloing G, Granadel C, Claverys JP. 2000. Cross-regulation of competence pheromone production and export in the early control of transformation in Streptococcus pneumoniae. Mol Microbiol 38:867-78.
93. Peterson S, Cline RT, Tettelin H, Sharov V, Morrison DA. 2000. Gene expression analysis of the Streptococcus pneumoniae competence regulons by use of DNA microarrays. J Bacteriol 182:6192-202.
94. Tovpeko Y, Bai J, Morrison DA. 2016. Competence for Genetic Transformation in Streptococcus pneumoniae: Mutations in σA Bypass the ComW Requirement for Late Gene Expression. J Bacteriol 198:2370-8.
95. Lee MS, Morrison DA. 1999. Identification of a new regulator in Streptococcus pneumoniae linking quorum sensing to competence for genetic transformation. J Bacteriol 181:5004-16.
96. Luo P, Morrison DA. 2003. Transient association of an alternative sigma factor, ComX, with RNA polymerase during the period of competence for genetic transformation in Streptococcus pneumoniae. J Bacteriol 185:349-58.
97. Peterson SN, Sung CK, Cline R, Desai BV, Snesrud EC, Luo P, Walling J, Li H, Mintz M, Tsegaye G, Burr PC, Do Y, Ahn S, Gilbert J, Fleischmann RD, Morrison DA. 2004. Identification of competence pheromone responsive genes in Streptococcus pneumoniae by use of DNA microarrays. Mol Microbiol 51:1051-70.
98. Javor GT, Tomasz A. 1968. An autoradiographic study of genetic transformation. Proc Natl Acad Sci U S A 60:1216-22.
99. Slager J, Aprianto R, Veening JW. 2019. Refining the pneumococcal competence regulon by RNA-sequencing. J Bacteriol.
100. Whatmore AM, Barcus VA, Dowson CG. 1999. Genetic diversity of the streptococcal competence (com) gene locus. J Bacteriol 181:3144-54.
101. Iannelli F, Oggioni MR, Pozzi G. 2005. Sensor domain of histidine kinase ComD confers competence pherotype specificity in Streptoccoccus pneumoniae. FEMS Microbiol Lett 252:321-6.
102. Li G, Liang Z, Wang X, Yang Y, Shao Z, Li M, Ma Y, Qu F, Morrison DA, Zhang JR. 2016. Addiction of Hypertransformable Pneumococcal Isolates to Natural Transformation for In Vivo Fitness and Virulence. Infect Immun 84:1887-1901.
103. Prudhomme M, Berge M, Martin B, Polard P. 2016. Pneumococcal Competence Coordination Relies on a Cell-Contact Sensing Mechanism. PLoS Genet 12:e1006113.
104. Moreno-Gámez S, Sorg RA, Domenech A, Kjos M, Weissing FJ, van Doorn GS, Veening JW. 2017. Quorum sensing integrates environmental cues, cell density and cell history to control bacterial competence. Nat Commun 8:854.
105. Marks LR, Reddinger RM, Hakansson AP. 2012. High levels of genetic recombination during nasopharyngeal carriage and biofilm formation in Streptococcus pneumoniae. MBio 3.
106. Tomasz A. 1966. Model for the mechanism controlling the expression of competent state in Pneumococcus cultures. J Bacteriol 91:1050-61.

107. Chen JD, Morrison DA. 1987. Modulation of competence for genetic transformation in Streptococcus pneumoniae. J Gen Microbiol 133:1959-67.
108. Slager J, Kjos M, Attaiech L, Veening JW. 2014. Antibiotic-induced replication stress triggers bacterial competence by increasing gene dosage near the origin. Cell 157:395-406.
109. Domenech A, Slager J, Veening JW. 2018. Antibiotic-Induced Cell Chaining Triggers Pneumococcal Competence by Reshaping Quorum Sensing to Autocrine-Like Signaling. Cell Rep 25:2390-2400.e3.
110. Stevens KE, Chang D, Zwack EE, Sebert ME. 2011. Competence in *Streptococcus pneumoniae* is regulated by the rate of ribosomal decoding errors. MBio 2.
111. Weyder M, Prudhomme M, Bergé M, Polard P, Fichant G. 2018. Dynamic Modeling of S*treptococcus pneumoniae* Competence Provides Regulatory Mechanistic Insights Into Its Tight Temporal Regulation. Front Microbiol 9:1637.
112. Cassone M, Gagne AL, Spruce LA, Seeholzer SH, Sebert ME. 2012. The HtrA protease from Streptococcus pneumoniae digests both denatured proteins and the competence-stimulating peptide. J Biol Chem 287:38449-59.
113. Laux A, Sexauer A, Sivaselvarajah D, Kaysen A, Brückner R. 2015. Control of competence by related non-coding csRNAs in Streptococcus pneumoniae R6. Front Genet 6:246.
114. Mascher T, Zähner D, Merai M, Balmelle N, de Saizieu AB, Hakenbeck R. 2003. The Streptococcus pneumoniae cia regulon: CiaR target sites and transcription profile analysis. J Bacteriol 185:60-70.
115. Dagkessamanskaia A, Moscoso M, Hénard V, Guiral S, Overweg K, Reuter M, Martin B, Wells J, Claverys JP. 2004. Interconnection of competence, stress and CiaR regulons in Streptococcus pneumoniae: competence triggers stationary phase autolysis of ciaR mutant cells. Mol Microbiol 51:1071-86.
116. Rimini R, Jansson B, Feger G, Roberts TC, de Francesco M, Gozzi A, Faggioni F, Domenici E, Wallace DM, Frandsen N, Polissi A. 2000. Global analysis of transcription kinetics during competence development in Streptococcus pneumoniae using high density DNA arrays. Mol Microbiol 36:1279-92.
117. Morrison DA. 1997. Streptococcal competence for genetic transformation: regulation by peptide pheromones. Microb Drug Resist 3:27-37.
118. Mirouze N, Bergé MA, Soulet AL, Mortier-Barrière I, Quentin Y, Fichant G, Granadel C, Noirot-Gros MF, Noirot P, Polard P, Martin B, Claverys JP. 2013. Direct involvement of DprA, the transformation-dedicated RecA loader, in the shut-off of pneumococcal competence. Proc Natl Acad Sci U S A 110:E1035-44.
119. Desai BV, Morrison DA. 2007. Transformation in Streptococcus pneumoniae: formation of eclipse complex in a coiA mutant implicates CoiA in genetic recombination. Mol Microbiol 63:1107-17.
120. Desai BV, Morrison DA. 2006. An unstable competence-induced protein, CoiA, promotes processing of donor DNA after uptake during genetic transformation in Streptococcus pneumoniae. J Bacteriol 188:5177-86.
121. Sanchez D, Boudes M, van Tilbeurgh H, Durand D, Quevillon-Cheruel S. 2015. Modeling the ComD/ComE/comcde interaction network using small angle X-ray scattering. FEBS J 282:1538-53.

122. Luo P, Li H, Morrison DA. 2004. Identification of ComW as a new component in the regulation of genetic transformation in Streptococcus pneumoniae. Mol Microbiol 54:172-83.
123. Piotrowski A, Luo P, Morrison DA. 2009. Competence for genetic transformation in Streptococcus pneumoniae: termination of activity of the alternative sigma factor ComX is independent of proteolysis of ComX and ComW. J Bacteriol 191:3359-66.
124. Johnsborg O, Eldholm V, Bjørnstad ML, Håvarstein LS. 2008. A predatory mechanism dramatically increases the efficiency of lateral gene transfer in Streptococcus pneumoniae and related commensal species. Mol Microbiol 69:245-53.
125. Eldholm V, Johnsborg O, Straume D, Ohnstad HS, Berg KH, Hermoso JA, Håvarstein LS. 2010. Pneumococcal CbpD is a murein hydrolase that requires a dual cell envelope binding specificity to kill target cells during fratricide. Mol Microbiol 76:905-17.
126. Mosser JL, Tomasz A. 1970. Choline-containing teichoic acid as a structural component of pneumococcal cell wall and its role in sensitivity to lysis by an autolytic enzyme. J Biol Chem 245:287-98.
127. Guiral S, Mitchell TJ, Martin B, Claverys JP. 2005. Competence-programmed predation of noncompetent cells in the human pathogen Streptococcus pneumoniae: genetic requirements. Proc Natl Acad Sci U S A 102:8710-5.
128. García P, Paz González M, García E, García JL, López R. 1999. The molecular characterization of the first autolytic lysozyme of Streptococcus pneumoniae reveals evolutionary mobile domains. Mol Microbiol 33:128-38.
129. Eldholm V, Johnsborg O, Haugen K, Ohnstad HS, Håvarstein LS. 2009. Fratricide in Streptococcus pneumoniae: contributions and role of the cell wall hydrolases CbpD, LytA and LytC. Microbiology 155:2223-34.
130. Håvarstein LS, Martin B, Johnsborg O, Granadel C, Claverys JP. 2006. New insights into the pneumococcal fratricide: relationship to clumping and identification of a novel immunity factor. Mol Microbiol 59:1297-307.
131. Wei H, Håvarstein LS. 2012. Fratricide is essential for efficient gene transfer between pneumococci in biofilms. Appl Environ Microbiol 78:5897-905.
132. Martin B, Granadel C, Campo N, Hénard V, Prudhomme M, Claverys JP. 2010. Expression and maintenance of ComD-ComE, the two-component signal-transduction system that controls competence of Streptococcus pneumoniae. Mol Microbiol 75:1513-28.
133. Miller EL, Evans BA, Cornejo OE, Roberts IS, Rozen DE. 2017. Pherotype Polymorphism in Streptococcus pneumoniae Has No Obvious Effects on Population Structure and Recombination. Genome Biol Evol 9:2546-2559.
134. Kilian M, Poulsen K, Blomqvist T, Håvarstein LS, Bek-Thomsen M, Tettelin H, Sørensen UB. 2008. Evolution of Streptococcus pneumoniae and its close commensal relatives. PLoS One 3:e2683.
135. Pinchas MD, LaCross NC, Dawid S. 2015. An electrostatic interaction between BlpC and BlpH dictates pheromone specificity in the control of bacteriocin production and immunity in Streptococcus pneumoniae. J Bacteriol 197:1236-48.

136. Miller EL, Kjos M, Abrudan MI, Roberts IS, Veening JW, Rozen DE. 2018. Eavesdropping and crosstalk between secreted quorum sensing peptide signals that regulate bacteriocin production in Streptococcus pneumoniae. ISME J 12:2363-2375.
137. Lux T, Nuhn M, Hakenbeck R, Reichmann P. 2007. Diversity of bacteriocins and activity spectrum in Streptococcus pneumoniae. J Bacteriol 189:7741-51.
138. Miller EL, Abrudan MI, Roberts IS, Rozen DE. 2016. Diverse Ecological Strategies Are Encoded by Streptococcus pneumoniae Bacteriocin-Like Peptides. Genome Biol Evol 8:1072-90.
139. Prudhomme M, Attaiech L, Sanchez G, Martin B, Claverys JP. 2006. Antibiotic stress induces genetic transformability in the human pathogen Streptococcus pneumoniae. Science 313:89-92.
140. Gagne AL, Stevens KE, Cassone M, Pujari A, Abiola OE, Chang DJ, Sebert ME. 2013. Competence in Streptococcus pneumoniae is a response to an increasing mutational burden. PLoS One 8:e72613.
141. Engelmoer DJ, Donaldson I, Rozen DE. 2013. Conservative sex and the benefits of transformation in Streptococcus pneumoniae. PLoS Pathog 9:e1003758.
142. Hiller NL, Ahmed A, Powell E, Martin DP, Eutsey R, Earl J, Janto B, Boissy RJ, Hogg J, Barbadora K, Sampath R, Lonergan S, Post JC, Hu FZ, Ehrlich GD. 2010. Generation of genic diversity among Streptococcus pneumoniae strains via horizontal gene transfer during a chronic polyclonal pediatric infection. PLoS Pathog 6:e1001108.
143. Chewapreecha C, Harris SR, Croucher NJ, Turner C, Marttinen P, Cheng L, Pessia A, Aanensen DM, Mather AE, Page AJ, Salter SJ, Harris D, Nosten F, Goldblatt D, Corander J, Parkhill J, Turner P, Bentley SD. 2014. Dense genomic sampling identifies highways of pneumococcal recombination. Nat Genet 46:305-309.
144. Chalkley L, Schuster C, Potgieter E, Hakenbeck R. 1991. Relatedness between Streptococcus pneumoniae and viridans streptococci: transfer of penicillin resistance determinants and immunological similarities of penicillin-binding proteins. FEMS Microbiol Lett 69:35-42.
145. Janoir C, Podglajen I, Kitzis MD, Poyart C, Gutmann L. 1999. In vitro exchange of fluoroquinolone resistance determinants between Streptococcus pneumoniae and viridans streptococci and genomic organization of the parE-parC region in S. mitis. J Infect Dis 180:555-8.
146. Cerdá Zolezzi P, Laplana LM, Calvo CR, Cepero PG, Erazo MC, Gómez-Lus R. 2004. Molecular basis of resistance to macrolides and other antibiotics in commensal viridans group streptococci and Gemella spp. and transfer of resistance genes to Streptococcus pneumoniae. Antimicrob Agents Chemother 48:3462-7.
147. Donati C, Hiller NL, Tettelin H, Muzzi A, Croucher NJ, Angiuoli SV, Oggioni M, Dunning Hotopp JC, Hu FZ, Riley DR, Covacci A, Mitchell TJ, Bentley SD, Kilian M, Ehrlich GD, Rappuoli R, Moxon ER, Masignani V. 2010. Structure and dynamics of the pan-genome of Streptococcus pneumoniae and closely related species. Genome Biol 11:R107.

148. Jensen A, Valdórsson O, Frimodt-Møller N, Hollingshead S, Kilian M. 2015. Commensal streptococci serve as a reservoir for β-lactam resistance genes in Streptococcus pneumoniae. Antimicrob Agents Chemother 59:3529-40.
149. Dowson CG, Coffey TJ, Kell C, Whiley RA. 1993. Evolution of penicillin resistance in Streptococcus pneumoniae; the role of Streptococcus mitis in the formation of a low affinity PBP2B in S. pneumoniae. Mol Microbiol 9:635-43.
150. Dowson CG, Hutchison A, Brannigan JA, George RC, Hansman D, Liñares J, Tomasz A, Smith JM, Spratt BG. 1989. Horizontal transfer of penicillin-binding protein genes in penicillin-resistant clinical isolates of Streptococcus pneumoniae. Proc Natl Acad Sci U S A 86:8842-6.
151. Coffey TJ, Dowson CG, Daniels M, Zhou J, Martin C, Spratt BG, Musser JM. 1991. Horizontal transfer of multiple penicillin-binding protein genes, and capsular biosynthetic genes, in natural populations of Streptococcus pneumoniae. Mol Microbiol 5:2255-60.
152. Ding F, Tang P, Hsu MH, Cui P, Hu S, Yu J, Chiu CH. 2009. Genome evolution driven by host adaptations results in a more virulent and antimicrobial-resistant Streptococcus pneumoniae serotype 14. BMC Genomics 10:158.
153. Bentley SD, Aanensen DM, Mavroidi A, Saunders D, Rabbinowitsch E, Collins M, Donohoe K, Harris D, Murphy L, Quail MA, Samuel G, Skovsted IC, Kaltoft MS, Barrell B, Reeves PR, Parkhill J, Spratt BG. 2006. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. PLoS Genet 2:e31.
154. Brueggemann AB, Pai R, Crook DW, Beall B. 2007. Vaccine escape recombinants emerge after pneumococcal vaccination in the United States. PLoS Pathog 3:e168.
155. Richter SS, Diekema DJ, Heilmann KP, Dohrn CL, Riahi F, Doern GV. 2014. Changes in pneumococcal serotypes and antimicrobial resistance after introduction of the 13-valent conjugate vaccine in the United States. Antimicrob Agents Chemother 58:6484-9.
156. Rezaei Javan R, van Tonder AJ, King JP, Harrold CL, Brueggemann AB. 2018. Genome Sequencing Reveals a Large and Diverse Repertoire of Antimicrobial Peptides. Front Microbiol 9:2012.
157. McAuliffe O, Ross RP, Hill C. 2001. Lantibiotics: structure, biosynthesis and mode of action. FEMS Microbiol Rev 25:285-308.
158. Martínez B, Suárez JE, Rodríguez A. 1996. Lactococcin 972 : a homodimeric lactococcal bacteriocin whose primary target is not the plasma membrane. Microbiology 142 (Pt 9):2393-8.
159. Gabrielsen C, Brede DA, Nes IF, Diep DB. 2014. Circular bacteriocins: biosynthesis and mode of action. Appl Environ Microbiol 80:6854-62.
160. Hegemann JD, Zimmermann M, Xie X, Marahiel MA. 2015. Lasso peptides: an intriguing class of bacterial natural products. Acc Chem Res 48:1909-19.
161. Flühe L, Marahiel MA. 2013. Radical S-adenosylmethionine enzyme catalyzed thioether bond formation in sactipeptide biosynthesis. Curr Opin Chem Biol 17:605-12.

162. Nissen-Meyer J, Oppegård C, Rogne P, Haugen HS, Kristiansen PE. 2010. Structure and Mode-of-Action of the Two-Peptide (Class-IIb) Bacteriocins. Probiotics Antimicrob Proteins 2:52-60.
163. Shen P, Lees JA, Bee GCW, Brown SP, Weiser JN. 2019. Pneumococcal quorum sensing drives an asymmetric owner-intruder competitive strategy during carriage via the competence regulon. Nat Microbiol 4:198-208.
164. Islam MR, Nagao J, Zendo T, Sonomoto K. 2012. Antimicrobial mechanism of lantibiotics. Biochem Soc Trans 40:1528-33.
165. Zhang Q, Yu Y, Vélasquez JE, van der Donk WA. 2012. Evolution of lanthipeptide synthetases. Proc Natl Acad Sci U S A 109:18361-6.
166. Asaduzzaman SM, Sonomoto K. 2009. Lantibiotics: diverse activities and unique modes of action. J Biosci Bioeng 107:475-87.
167. Majchrzykiewicz JA, Lubelski J, Moll GN, Kuipers A, Bijlsma JJ, Kuipers OP, Rink R. 2010. Production of a class II two-component lantibiotic of Streptococcus pneumoniae using the class I nisin synthetic machinery and leader sequence. Antimicrob Agents Chemother 54:1498-505.
168. Hoover SE, Perez AJ, Tsui HC, Sinha D, Smiley DL, DiMarchi RD, Winkler ME, Lazazzera BA. 2015. A new quorum-sensing system (TprA/PhrA) for Streptococcus pneumoniae D39 that regulates a lantibiotic biosynthesis gene cluster. Mol Microbiol 97:229-43.
169. Kadam A, Eutsey RA, Rosch J, Miao X, Longwell M, Xu W, Woolford CA, Hillman T, Motib AS, Yesilkaya H, Mitchell AP, Hiller NL. 2017. Promiscuous signaling by a regulatory system unique to the pandemic PMEN1 pneumococcal lineage. PLoS Pathog 13:e1006339.
170. Maricic N, Anderson ES, Opipari AE, Yu EA, Dawid S. 2016. Characterization of a Multipeptide Lantibiotic Locus in Streptococcus pneumoniae. MBio 7:e01656-15.
171. Dawid S, Roche AM, Weiser JN. 2007. The blp bacteriocins of Streptococcus pneumoniae mediate intraspecies competition both in vitro and in vivo. Infect Immun 75:443-51.
172. Bogaardt C, van Tonder AJ, Brueggemann AB. 2015. Genomic analyses of pneumococci reveal a wide diversity of bacteriocins - including pneumocyclicin, a novel circular bacteriocin. BMC Genomics 16:554.
173. Wholey WY, Abu-Khdeir M, Yu EA, Siddiqui S, Esimai O, Dawid S. 2019. Characterization of the Competitive Pneumocin Peptides of *Streptococcus pneumoniae*. Front Cell Infect Microbiol 9:55.
174. Valente C, Dawid S, Pinto FR, Hinds J, Simões AS, Gould KA, Mendes LA, de Lencastre H, Sá-Leão R. 2016. The blp Locus of Streptococcus pneumoniae Plays a Limited Role in the Selection of Strains That Can Cocolonize the Human Nasopharynx. Appl Environ Microbiol 82:5206-15.
175. Aniansson G, Alm B, Andersson B, Larsson P, Nylen O, Peterson H, Rigner P, Svanborg M, Svanborg C. 1992. Nasopharyngeal colonization during the first year of life. J Infect Dis 165 Suppl 1:S38-42.
176. Kononen E, Jousimies-Somer H, Bryk A, Kilp T, Kilian M. 2002. Establishment of streptococci in the upper respiratory tract: longitudinal changes in the mouth and nasopharynx up to 2 years of age. J Med Microbiol 51:723-30.

177. Bogaert D, Keijser B, Huse S, Rossen J, Veenhoven R, van Gils E, Bruin J, Montijn R, Bonten M, Sanders E. 2011. Variability and diversity of nasopharyngeal microbiota in children: a metagenomic analysis. PLoS One 6:e17035.
178. Luo P, Li H, Morrison DA. 2003. ComX is a unique link between multiple quorum sensing outputs and competence in Streptococcus pneumoniae. Mol Microbiol 50:623-33.
179. de Saizieu A, Gardes C, Flint N, Wagner C, Kamber M, Mitchell TJ, Keck W, Amrein KE, Lange R. 2000. Microarray-based identification of a novel *Streptococcus pneumoniae* regulon controlled by an autoinduced peptide. J Bacteriol 182:4696-703.
180. Claverys JP, Martin B, Havarstein LS. 2007. Competence-induced fratricide in streptococci. Mol Microbiol 64:1423-33.
181. Hall MP, Unch J, Binkowski BF, Valley MP, Butler BL, Wood MG, Otto P, Zimmerman K, Vidugiris G, Machleidt T, Robers MB, Benink HA, Eggers CT, Slater MR, Meisenheimer PL, Klaubert DH, Fan F, Encell LP, Wood KV. 2012. Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. ACS Chem Biol 7:1848-57.
182. Michelini E, Cevenini L, Mezzanotte L, Ablamsky D, Southworth T, Branchini B, Roda A. 2008. Spectral-resolved gene technology for multiplexed bioluminescence and high-content screening. Anal Chem 80:260-7.
183. Kochan TJ, Dawid S. 2013. The HtrA protease of Streptococcus pneumoniae controls density-dependent stimulation of the bacteriocin blp locus via disruption of pheromone secretion. J Bacteriol 195:1561-72.
184. Dixon AS, Schwinn MK, Hall MP, Zimmerman K, Otto P, Lubben TH, Butler BL, Binkowski BF, Machleidt T, Kirkland TA, Wood MG, Eggers CT, Encell LP, Wood KV. 2016. NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. ACS Chem Biol 11:400-8.
185. Schwinn MK, Machleidt T, Zimmerman K, Eggers CT, Dixon AS, Hurst R, Hall MP, Encell LP, Binkowski BF, Wood KV. 2017. CRISPR-Mediated Tagging of Endogenous Proteins with a Luminescent Peptide. ACS Chem Biol.
186. Ishii S, Fukui K, Yokoshima S, Kumagai K, Beniyama Y, Kodama T, Fukuyama T, Okabe T, Nagano T, Kojima H, Yano T. 2017. High-throughput Screening of Small Molecule Inhibitors of the Streptococcus Quorum-sensing Signal Pathway. Sci Rep 7:4029.
187. Dawid S, Sebert ME, Weiser JN. 2009. Bacteriocin activity of *Streptococcus pneumoniae* is controlled by the serine protease HtrA via posttranscriptional regulation. J Bacteriol 191:1509-18.
188. Sung CK, Li H, Claverys JP, Morrison DA. 2001. An rpsL cassette, janus, for gene replacement through negative selection in Streptococcus pneumoniae. Appl Environ Microbiol 67:5190-6.
189. Li Y, Thompson CM, Lipsitch M. 2014. A modified Janus cassette (Sweet Janus) to improve allelic replacement efficiency by high-stringency negative selection in Streptococcus pneumoniae. PLoS One 9:e100510.
190. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3--new capabilities and interfaces. Nucleic Acids Res 40:e115.

191. Koressaar T, Remm M. 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23:1289-91.
192. Puigbò P, Guzmán E, Romeu A, Garcia-Vallvé S. 2007. OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. Nucleic Acids Res 35:W126-31.
193. Wayne KJ, Sham LT, Tsui HC, Gutu AD, Barendt SM, Keen SK, Winkler ME. 2010. Localization and cellular amounts of the WalRKJ (VicRKX) two-component regulatory system proteins in serotype 2 Streptococcus pneumoniae. J Bacteriol 192:4388-94.
194. Schmidt TG, Batz L, Bonet L, Carl U, Holzapfel G, Kiem K, Matulewicz K, Niermeier D, Schuchardt I, Stanar K. 2013. Development of the Twin-Strep-tag® and its application for purification of recombinant proteins from cell culture supernatants. Protein Expr Purif 92:54-61.
195. Davey L, Halperin SA, Lee SF. 2015. Immunoblotting conditions for small peptides from streptococci. J Microbiol Methods 114:40-2.
196. Eran Y, Getter Y, Baruch M, Belotserkovsky I, Padalon G, Mishalian I, Podbielski A, Kreikemeyer B, Hanski E. 2007. Transcriptional regulation of the sil locus by the SilCR signalling peptide and its implications on group A streptococcus virulence. Mol Microbiol 63:1209-22.
197. Qi F, Chen P, Caufield PW. 1999. Functional analyses of the promoters in the lantibiotic mutacin II biosynthetic locus in Streptococcus mutans. Appl Environ Microbiol 65:652-8.
198. Fleuchot B, Gitton C, Guillot A, Vidic J, Nicolas P, Besset C, Fontaine L, Hols P, Leblond-Bourget N, Monnet V, Gardan R. 2011. Rgg proteins associated with internalized small hydrophobic peptides: a new quorum-sensing mechanism in streptococci. Mol Microbiol 80:1102-19.
199. Pérez-Pascual D, Gaudu P, Fleuchot B, Besset C, Rosinski-Chupin I, Guillot A, Monnet V, Gardan R. 2015. RovS and its associated signaling peptide form a cell-to-cell communication system required for Streptococcus agalactiae pathogenesis. MBio 6.
200. Chang JC, Federle MJ. 2016. PptAB Exports Rgg Quorum-Sensing Peptides in Streptococcus. PLoS One 11:e0168461.
201. Dmitriev AV, McDowell EJ, Kappeler KV, Chaussee MA, Rieck LD, Chaussee MS. 2006. The Rgg regulator of Streptococcus pyogenes influences utilization of nonglucose carbohydrates, prophage induction, and expression of the NAD-glycohydrolase virulence operon. J Bacteriol 188:7230-41.
202. Zhi X, Abdullah IT, Gazioglu O, Manzoor I, Shafeeq S, Kuipers OP, Hiller NL, Andrew PW, Yesilkaya H. 2018. Rgg-Shp regulators are important for pneumococcal colonization and invasion through their effect on mannose utilization and capsule synthesis. Sci Rep 8:6369.
203. Junges R, Salvadori G, Shekhar S, Åmdal HA, Periselneris JN, Chen T, Brown JS, Petersen FC. 2017. A Quorum-Sensing System That Regulates *Streptococcus pneumoniae* Biofilm Formation and Surface Polysaccharide Production. mSphere 2.
204. Fontaine L, Boutry C, de Frahan MH, Delplace B, Fremaux C, Horvath P, Boyaval P, Hols P. 2010. A novel pheromone quorum-sensing system controls

the development of natural competence in Streptococcus thermophilus and Streptococcus salivarius. J Bacteriol 192:1444-54.
205. Mashburn-Warren L, Morrison DA, Federle MJ. 2010. A novel double-tryptophan peptide pheromone controls competence in Streptococcus spp. via an Rgg regulator. Mol Microbiol 78:589-606.
206. Wang CY, Patel N, Wholey WY, Dawid S. 2018. ABC transporter content diversity in *Streptococcus pneumoniae* impacts competence regulation and bacteriocin production. Proc Natl Acad Sci U S A 115:E5776-E5785.
207. Hiller NL, Janto B, Hogg JS, Boissy R, Yu S, Powell E, Keefe R, Ehrlich NE, Shen K, Hayes J, Barbadora K, Klimke W, Dernovoy D, Tatusova T, Parkhill J, Bentley SD, Post JC, Ehrlich GD, Hu FZ. 2007. Comparative genomic analyses of seventeen Streptococcus pneumoniae strains: insights into the pneumococcal supragenome. J Bacteriol 189:8186-95.
208. Piet JR, Geldhoff M, van Schaik BD, Brouwer MC, Valls Seron M, Jakobs ME, Schipper K, Pannekoek Y, Zwinderman AH, van der Poll T, van Kampen AH, Baas F, van der Ende A, van de Beek D. 2014. Streptococcus pneumoniae arginine synthesis genes promote growth and virulence in pneumococcal meningitis. J Infect Dis 209:1781-91.
209. Agaisse H, Gominet M, Okstad OA, Kolstø AB, Lereclus D. 1999. PlcR is a pleiotropic regulator of extracellular virulence factor gene expression in Bacillus thuringiensis. Mol Microbiol 32:1043-53.
210. Bae T, Kozlowicz B, Dunny GM. 2002. Two targets in pCF10 DNA for PrgX binding: their role in production of Qa and prgX mRNA and in regulation of pheromone-inducible conjugation. J Mol Biol 315:995-1007.
211. Lasarre B, Aggarwal C, Federle MJ. 2013. Antagonistic Rgg regulators mediate quorum sensing via competitive DNA binding in Streptococcus pyogenes. MBio 3.
212. Croucher NJ, Finkelstein JA, Pelton SI, Mitchell PK, Lee GM, Parkhill J, Bentley SD, Hanage WP, Lipsitch M. 2013. Population genomics of post-vaccine changes in pneumococcal epidemiology. Nat Genet 45:656-63.
213. Gardan R, Besset C, Gitton C, Guillot A, Fontaine L, Hols P, Monnet V. 2013. Extracellular life cycle of ComS, the competence-stimulating peptide of Streptococcus thermophilus. J Bacteriol 195:1845-55.
214. Underhill SAM, Shields RC, Kaspar JR, Haider M, Burne RA, Hagen SJ. 2018. Intracellular Signaling by the *comRS* System in *Streptococcus mutans* Genetic Competence. mSphere 3.
215. Aprianto R, Slager J, Holsappel S, Veening JW. 2018. High-resolution analysis of the pneumococcal transcriptome under a wide range of infection-relevant conditions. Nucleic Acids Res 46:9990-10006.
216. Nes IF, Diep DB, Håvarstein LS, Brurberg MB, Eijsink V, Holo H. 1996. Biosynthesis of bacteriocins in lactic acid bacteria. Antonie Van Leeuwenhoek 70:113-28.
217. Jolley KA, Maiden MC. 2010. BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 11:595.
218. Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23:254-67.

219. Maricic N, Dawid S. 2014. Using the overlay assay to qualitatively measure bacterial production of and sensitivity to pneumococcal bacteriocins. J Vis Exp:e51876.
220. Overkamp W, Beilharz K, Detert Oude Weme R, Solopova A, Karsens H, Kovács Á, Kok J, Kuipers OP, Veening JW. 2013. Benchmarking various green fluorescent protein variants in Bacillus subtilis, Streptococcus pneumoniae, and Lactococcus lactis for live cell imaging. Appl Environ Microbiol 79:6481-90.
221. Liu X, Gallay C, Kjos M, Domenech A, Slager J, van Kessel SP, Knoops K, Sorg RA, Zhang JR, Veening JW. 2017. High-throughput CRISPRi phenotyping identifies new essential genes in *Streptococcus pneumoniae*. Mol Syst Biol 13:931.
222. Guiral S, Hénard V, Laaberki MH, Granadel C, Prudhomme M, Martin B, Claverys JP. 2006. Construction and evaluation of a chromosomal expression platform (CEP) for ectopic, maltose-driven gene expression in Streptococcus pneumoniae. Microbiology 152:343-9.
223. Chen YJ, Liu P, Nielsen AA, Brophy JA, Clancy K, Peterson T, Voigt CA. 2013. Characterization of 582 natural and synthetic terminators and quantification of their design constraints. Nat Methods 10:659-64.