



# Molecular Determinants of Substrate Selectivity of a Pneumococcal Rgg-Regulated Peptidase-Containing ABC Transporter

Charles Y. Wang,<sup>a</sup> Jennifer S. Medlin,<sup>b</sup> Don R. Nguyen,<sup>b\*</sup> W. Miguel Disbennett,<sup>b\*</sup>  Suzanne Dawid<sup>a,b</sup>

<sup>a</sup>Department of Microbiology and Immunology, University of Michigan, Ann Arbor, Michigan, USA

<sup>b</sup>Department of Pediatrics, University of Michigan, Ann Arbor, Michigan, USA

**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 coregulated *rtg* GG peptides. ComAB and BlpAB, which share a substrate pool, 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.

**IMPORTANCE** The export of peptides from the cell is a fundamental process carried out by all bacteria. One method of bacterial peptide export relies on a family of transporters called peptidase-containing ABC transporters (PCATs). PCATs export so-called GG peptides which carry out diverse functions, including cell-to-cell communication and interbacterial competition. In this work, we describe a PCAT-encoding genetic locus, *rtg*, in the pathogen *Streptococcus pneumoniae* (pneumococcus). The *rtg* locus is linked to increased competitive fitness advantage in a mouse model of nasopharyngeal colonization. We also describe how the *rtg* PCAT preferentially secretes a set of coregulated GG peptides but not GG peptides secreted by other pneumococcal PCATs. These findings illuminate a relatively understudied part of PCAT biology: how these transporters discriminate between different subsets of GG peptides. Ultimately, expanding our knowledge of PCATs will advance our understanding of the many microbial processes dependent on these transporters.

**KEYWORDS** ABC transporters, *Streptococcus pneumoniae*, bacteriocins, quorum sensing

**Citation** Wang CY, Medlin JS, Nguyen DR, Disbennett WM, Dawid S. 2020. Molecular determinants of substrate selectivity of a pneumococcal Rgg-regulated peptidase-containing ABC transporter. *mBio* 11:e02502-19. <https://doi.org/10.1128/mBio.02502-19>.

**Invited Editor** Jason W. Rosch, SJCRH

**Editor** Larry S. McDaniel, University of Mississippi Medical Center

**Copyright** © 2020 Wang et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Suzanne Dawid, [sdawid@umich.edu](mailto:sdawid@umich.edu).

\* Present address: Don R. Nguyen, Michigan State University, East Lansing, Michigan, USA; W. Miguel Disbennett, Ohio State University, Columbus, Ohio, USA.

**Received** 24 September 2019

**Accepted** 23 December 2019

**Published** 11 February 2020

**E**xport 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).

PCATs are ABC transporters that contain characteristic N-terminal peptidase domains (PEPs). PEP belongs to the family of C39 cysteine proteases and is responsible for the proteolytic processing of substrates during transport (3). In Gram-positive bacteria, PCATs function either alone or with a single additional accessory protein (4). 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 (5). Some PCATs also promote cell-to-cell communication by secreting the peptide pheromones of Gram-positive quorum-sensing systems (6–10). 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 (11). 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) (12). SHPs are small peptides which are exported by the PptAB transporter (13, 14) 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 (12). Besides PCATs and bacteriocin production, Rgg/SHP systems have been found to regulate diverse processes such as carbohydrate utilization (15), tissue invasion (13, 16), capsule production (16, 17), and biofilm formation (17, 18). 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 (19, 20).

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) (6, 7, 21, 22). ComAB and BlpAB secrete the same GG peptides, including the competence- and pneumocin-inducing pheromones and the pneumocins (23–25). Substrate sharing between ComAB and BlpAB affects competence and pneumocin regulation and influences when and with what effectiveness naturally occurring BlpAB<sup>+</sup> and BlpAB<sup>−</sup> strains can employ pneumocin-mediated killing (25, 26).

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 (3). 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 (27–29). 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 (28, 30, 31). 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 nonconserved residues to substrate-PEP binding (27–29).

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 (28, 32–35). These data suggest that in general, PCATs are not particularly selective when it comes to choosing substrates.

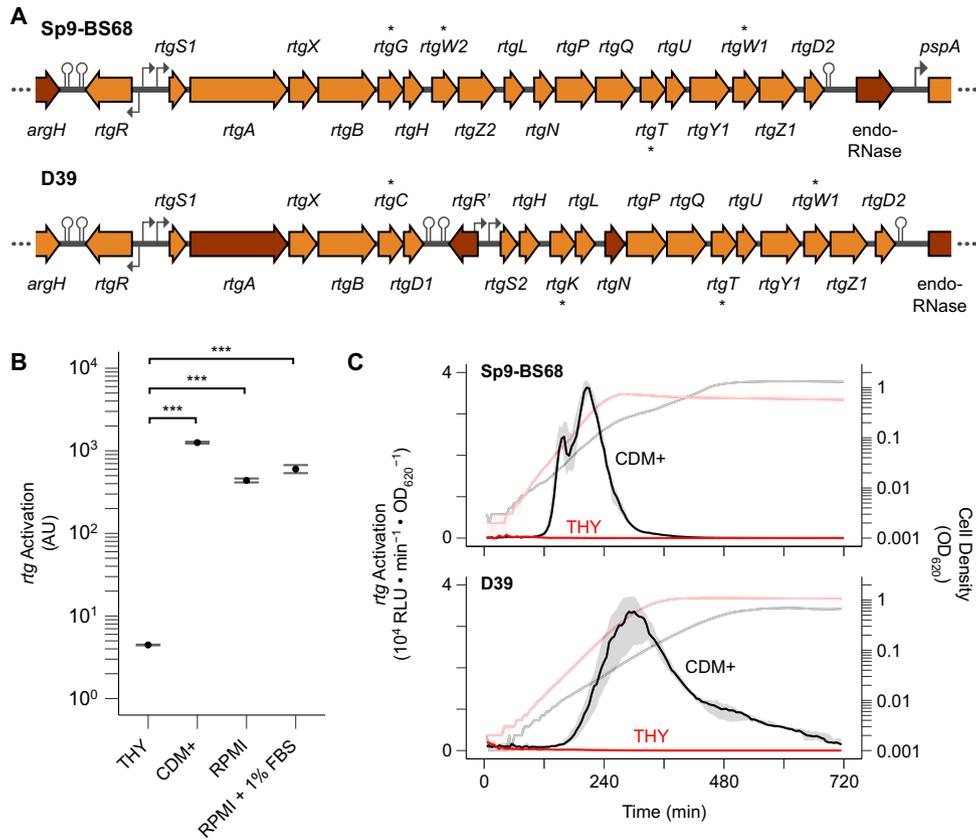
In this work, we describe a previously uncharacterized locus in pneumococcus, *rtg*, which encodes the PCAT RtgAB and several GG peptides. This locus is regulated by the Rgg/SHP-like system RtgR/S, which provides a competitive fitness advantage during nasopharyngeal colonization. We demonstrate that RtgAB secretes the *rtg* GG peptides but not ComAB/BlpAB substrates and that ComAB or BlpAB cannot 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.

## RESULTS

**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 (36) (Fig. 1A). Upstream of *07257* is a gene oriented in the opposite direction predicted to encode an Rgg-family transcription regulator (12). We hypothesized that this regulator controls expression of *07257* and named the locus *rtg* (Rgg-regulated transporter of double-glycine peptides). We designated 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 (37) 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 found in the laboratory strain D39 (and its derivative, R6) but with a disrupted *rtgA* (Fig. 1A).

**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 plus 0.5% yeast extract) but highly upregulated in the peptide-poor media CDM+ (38) and RPMI (Fig. 1B). The start of *rtg* activation in both Sp9-BS68 and D39 during growth in CDM+ occurs in early exponential phase at cell densities as low as an optical density at 620 nm ( $OD_{620}$ ) of 0.01 (Fig. 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 open reading frame (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. 1A).

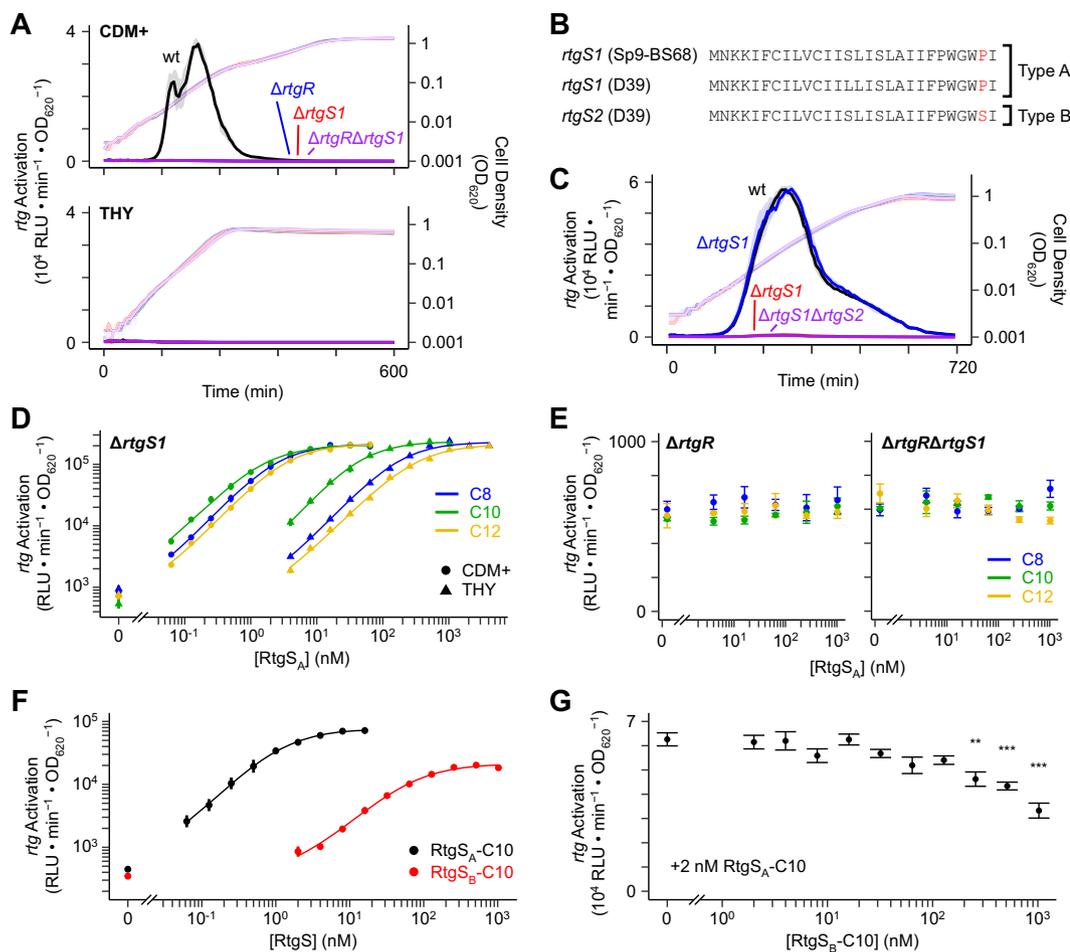
Having identified a putative Rgg/SHP-like regulatory system, we sought to define the contributions of RtgR and RtgS to *rtg* regulation through deletional analysis. We monitored *rtg* activation in Sp9-BS68  $\Delta$ *rtgS1*,  $\Delta$ *rtgR*, and  $\Delta$ *rtgS1*  $\Delta$ *rtgR* strains during growth in CDM+ and THY (Fig. 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*



**FIG 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 an  $OD_{620}$  of 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 to account for signal difference due to medium alone. Plots show means  $\pm$  standard errors (SEs) from 3 independent experiments. \*\*\*,  $P < 0.001$  by ANOVA with Tukey's honestly significant difference (HSD); AU, arbitrary units. (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 medians (lines) and 25% to 75% quantiles (shading) from 12 wells pooled from 3 independent experiments.

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. 2B). We found that the D39  $\Delta$ *rtgS1* and  $\Delta$ *rtgS1*  $\Delta$ *rtgS2* mutants failed to activate *rtg* in CDM+, while the  $\Delta$ *rtgS2* mutant was indistinguishable from the wild-type strain (Fig. 2C). This suggested 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<sub>A</sub>) and the *rtgS2* product in D39 as type B (RtgS<sub>B</sub>).

To confirm that RtgS is the specific pheromone inducer of *rtg*, we performed dose-response assays using synthetic peptides corresponding to the C-terminal 8, 10, and 12 residues of RtgS<sub>A</sub> (RtgS<sub>A</sub>-C8, RtgS<sub>A</sub>-C10, and RtgS<sub>A</sub>-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. 2D; Table 1). We also confirmed that *rtg* induction by synthetic RtgS requires RtgR (Fig. 2E). In D39, RtgS<sub>A</sub>-C10 induces *rtg* similarly to Sp9-BS68, whereas RtgS<sub>B</sub>-C10 acts as a partial agonist with a 55-fold larger 50% effective concentration ( $EC_{50}$ ) value than RtgS<sub>A</sub>-C10 (Fig. 2F; Table 1). Therefore,



**FIG 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 medians (lines) and 25% to 75% quantiles (shading) from 12 wells pooled from 3 independent experiments. (B) Translated *rtgS* gene products from Sp9-BS68 and D39. The type-defining residues are 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 medians (lines) and 25% to 75% quantiles (shading) from 30 wells pooled from 3 independent experiments. C-terminal fragments of RtgS<sub>A</sub> induce *rtg* in an RtgR-dependent manner. Sp9-BS68  $P_{rtgS1}$ -*luc* reporters were grown in CDM+ and THY (D) or CDM+ only (E) to an OD<sub>620</sub> of 0.02 and treated with synthetic RtgS<sub>A</sub> fragments. Response was defined as the maximum observed  $P_{rtgS1}$  activity within 60 min of treatment. Data in panel D were fitted to the four-parameter Hill model (curves). Plotted data points represent means  $\pm$  SEs from 3 independent experiments. (F) RtgS<sub>B</sub> is a partial agonist at the *rtg* locus. A D39  $\Delta rtgS1 \Delta rtgS2$   $P_{rtgS1}$ -*luc* reporter was grown in CDM+ to an OD<sub>620</sub> of 0.02 and treated with synthetic RtgS fragments. Response was defined as the maximum observed  $P_{rtgS1}$  activity within 60 min of treatment. Data were fitted to the four-parameter Hill model (curves). Plotted data points represent means  $\pm$  SEs from 3 independent experiments. (G) High concentrations of RtgS<sub>B</sub>-C10 antagonize RtgS<sub>A</sub>-C10. A D39  $\Delta rtgS1 \Delta rtgS2$   $P_{rtgS1}$ -*luc* reporter was grown in CDM+ to an OD<sub>620</sub> of 0.02 and treated simultaneously with 2 nM RtgS<sub>A</sub>-C10 and various concentrations of RtgS<sub>B</sub>-C10. Response was defined as the maximum observed  $P_{rtgS1}$  activity within 120 min of treatment. Plotted data points represent means  $\pm$  SEs from 3 independent experiments. \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$  versus 0 nM RtgS<sub>B</sub>-C10 by ANOVA with Dunnett's correction for multiple comparisons.

the Pro-to-Ser substitution in RtgS<sub>B</sub> 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<sub>B</sub> during natural *rtg* activation (Fig. 2C), and competitive dose-response assays showed that RtgS<sub>B</sub>-C10 only antagonizes RtgS<sub>A</sub>-C10 at likely supraphysiological concentrations ( $\geq 256$  nM) (Fig. 2G).

Next, we determined that, consistent with previously described Rgg/SHP systems, *rtg* activation requires both the Ami importer and the PptAB transporter (see Fig. S1A in the supplemental material). We also showed that response to exogenous RtgS treatment requires Ami but not PptAB (Fig. S1B), consistent with their respective roles as pheromone importer and exporter.

**TABLE 1** RtgS dose-response parameters

Strain	Growth medium	Peptide	EC <sub>50</sub> (nM) <sup>a</sup>	Maximal response (10 <sup>3</sup> RLU·min <sup>-1</sup> ·OD <sub>620</sub> <sup>-1</sup> ) <sup>a</sup>
Sp9-BS68 <sup>b</sup>	CDM+	RtgS <sub>A</sub> -C8	2.53 ± 0.25	212.2 ± 9.4
		RtgS <sub>A</sub> -C10	1.63 ± 0.19	205.7 ± 9.1
		RtgS <sub>A</sub> -C12	3.51 ± 0.28	213.9 ± 8.0
	THY	RtgS <sub>A</sub> -C8	184 ± 15	229.7 ± 8.6
		RtgS <sub>A</sub> -C10	47.4 ± 6.9	232.4 ± 14.3
D39 <sup>c</sup>	CDM+	RtgS <sub>A</sub> -C10	1.25 ± 0.30	75.1 ± 8.1
		RtgS <sub>B</sub> -C10	68.8 ± 11.5	21.2 ± 1.5

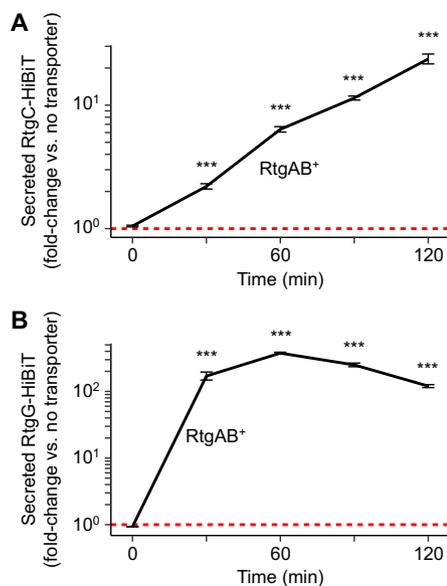
<sup>a</sup>Means ± SEs; derived from fitting data to Hill model.

<sup>b</sup> $\Delta$ rtgS1.

<sup>c</sup> $\Delta$ rtgS1  $\Delta$ rtgS2.

Finally, mutagenesis of the Sp9-BS68 *rtgS1* promoter region revealed the presence of two nearly identical promoters, each contributing partially to *rtgS1* expression (see Fig. S2A and B). We also identified an inverted repeat found in both promoters which is required for RtgS-induced expression and likely represents the RtgR binding site (Fig. S2A and C).

**RtgAB secretes *rtg*-encoded GG peptides.** Sp9-BS68 and D39 both harbor four putative GG peptides at the *rtg* locus (Fig. 1A). To determine whether RtgAB secretes the *rtg* GG peptides, we employed a HiBiT tag-based peptide secretion assay (25). We constructed autoinducing-deficient  $\Delta$ rtgS1  $\Delta$ rtgS2 R6 strains which harbor HiBiT-tagged *rtgC* (from D39) or *rtgG* (from Sp9-BS68) placed downstream of *rtgB* in RtgAB<sup>-</sup> (native R6 *rtgA* pseudogene) and RtgAB<sup>+</sup> (pseudogene-repaired) backgrounds. These strains were also ComAB<sup>-</sup> and BlpAB<sup>-</sup> in order to remove the possibility of peptide secretion through these other PCATs. Upon RtgS-C10 treatment in CDM+, levels of extracellular RtgC-HiBiT and RtgG-HiBiT in the RtgAB<sup>+</sup> cultures increased 26- and 376-fold, respectively, compared to their levels in the RtgAB<sup>-</sup> cultures (Fig. 3). From these data, we conclude that RtgAB secretes the *rtg* GG peptides.



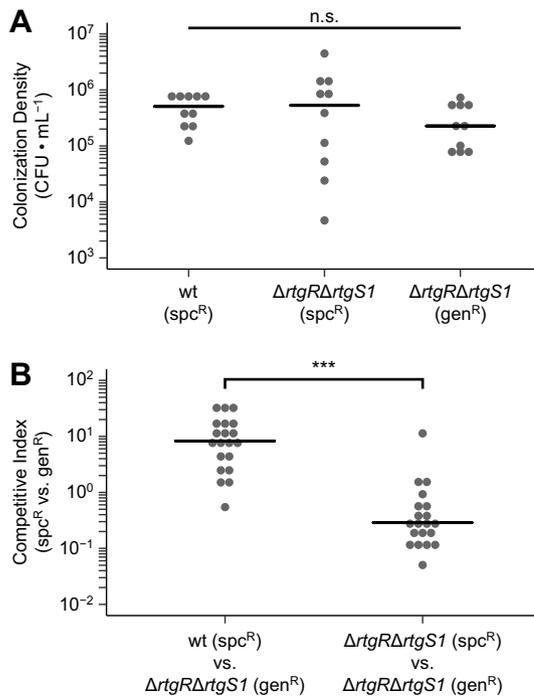
**FIG 3** RtgAB secretes RtgC and RtgG. R6 ComAB<sup>-</sup>/BlpAB<sup>-</sup>/RtgAB<sup>-</sup> and ComAB<sup>-</sup>/BlpAB<sup>-</sup>/RtgAB<sup>+</sup> strains expressing RtgC-HiBiT (A) or RtgG-HiBiT (B) were grown in CDM+ to an OD<sub>620</sub> of 0.05 and treated with 200 ng/ml CSP, 200 ng/ml BlpC, and 20 nM RtgS<sub>A</sub>-C10. CSP and BlpC were added to allow for comparison with ComAB<sup>+</sup> and BlpAB<sup>+</sup> strains in later experiments. Samples were taken every 30 min, and extracellular HiBiT signal was quantified. Data are presented as fold change values relative to the ComAB<sup>-</sup>/BlpAB<sup>-</sup>/RtgAB<sup>-</sup> control. Red dashed lines represent fold change = 1 (no difference versus the control). Plots show means ± SEs from 3 independent experiments. \*\*\*, *P* < 0.001 versus ComAB<sup>-</sup>/BlpAB<sup>-</sup>/RtgAB<sup>-</sup> by ANOVA with Tukey's HSD.

**TABLE 2** Genes of the *rtg* locus

Gene	Product	Frequency (%) <sup>a</sup>
<i>rtgA</i>		
Intact	Peptidase-containing ABC transporter	5.0
Start codon mutation only	Peptidase-containing ABC transporter	14.5
Disrupted	Peptidase-containing ABC transporter, truncated	59.1
<i>rtgB</i>		
Intact	Putative transport accessory protein	18.6
Disrupted	Putative transport accessory protein, truncated	2.5
<i>rtgC</i>	GG peptide	15.1
<i>rtgD</i>	Hypothetical protein	99.7
<i>rtgE</i>	GG peptide	7.2
<i>rtgG</i>	GG peptide	4.1
<i>rtgH</i>	Hypothetical protein	40.6
<i>rtgK</i>	GG peptide	41.2
<i>rtgL</i>	Hypothetical protein	83.3
<i>rtgM</i>	GG peptide	59.7
<i>rtgN</i>	Hypothetical protein	82.4
<i>rtgP</i>	Putative integral membrane protein	83.3
<i>rtgQ</i>	Putative integral membrane protein	83.3
<i>rtgR</i>	Rgg-family transcription regulator	78.6
<i>rtgR'</i>	Rgg-family transcription regulator, truncated	36.5
<i>rtgS</i>		
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
<i>rtgT</i>	GG peptide	83.0
<i>rtgU</i>	Hypothetical protein	83.0
<i>rtgV</i>	GG peptide; probable fusion of RtgW and RtgZ	35.5
<i>rtgW</i>	GG peptide	83.0
<i>rtgX</i>	Gyrl-like hypothetical protein	21.1
<i>rtgY</i>	Putative integral membrane protein	88.7
<i>rtgZ</i>	Putative integral membrane protein	99.7

<sup>a</sup>In 318 genomes from the Massachusetts collection with fully sequenced gapless *rtg* loci. May include pseudogenes unless stated otherwise.

**The *rtg* locus exhibits extensive variation across different pneumococcal strains.** To catalog the diversity found at *rtg*, we conducted a survey of the locus in a collection of pneumococcal clinical isolates from Massachusetts, USA (39). After removing genomes in which *rtg* spans multiple contigs or when conserved genes flanking *rtg* could not be found, we were left with 318 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 (see Fig. S3A). Across all 318 loci, we found 24 unique *rtg* genes (Table 2), including eight putative GG peptides which share highly conserved signal sequences (Fig. S3B). Searching for these genes in the full collection revealed that 615 of 616 strains had some version of the *rtg* locus. Next, we analyzed the variation in RtgS among the 318 filtered strains. 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 showed 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 that is otherwise intact. We determined

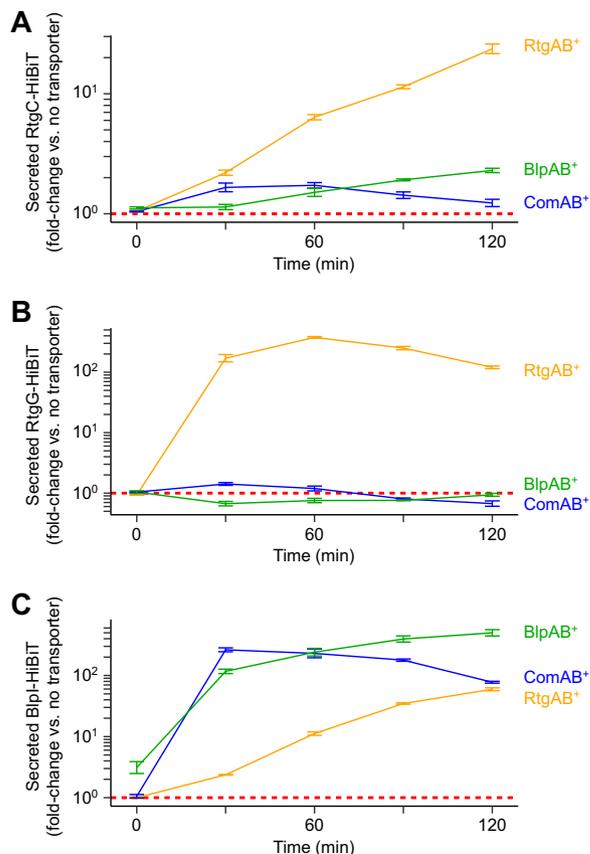


**FIG 4** Active RtgR/S provides a competitive fitness advantage during nasopharyngeal colonization. (A) Sp9-BS68 strains were singly inoculated into the nasopharynx of 5- to 7-week-old female BALB/c mice. At 3 days postinoculation, colonization density was sampled by nasal wash. Black bars represent medians.  $N = 10$  mice per strain pooled from 2 independent experiments. n.s., not significant by Kruskal-Wallis test. (B) Pairs of Sp9-BS68 strains were coinoculated into the nasopharynxes of 5- to 7-week-old female BALB/c mice. At 3 days postinoculation, 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. \*\*\*,  $P < 0.001$  by Mann-Whitney test.

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 that of a strain with fully intact *rtgA* (see Fig. S4). Therefore, 17% of strains encode a functional RtgAB.

**Active RtgR/S confers a competitive fitness advantage during nasopharyngeal colonization.** 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 wild-type and  $\Delta rtgR \Delta rtgS1$  strains in singly inoculated mice at 3 days postinoculation (Fig. 4A), the wild-type strain outcompeted the mutant in coinoculated mice (Fig. 4B). 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 cocolonization.

**RtgAB and ComAB/BlpAB preferentially secrete different sets of peptides.** The pneumococcal PCATs ComAB and BlpAB secrete the same diverse set of GG peptides (23–25). 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<sup>+</sup> and BlpAB<sup>+</sup> strains, using treatment with the *com* and *blp* pheromones competence-stimulating peptide (CSP) and BlpC, respectively, to induce their expression. ComAB and BlpAB secrete markedly reduced amounts of RtgC-HiBiT and RtgG-HiBiT compared to the amounts secreted by RtgAB (Fig. 5A and 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 10-fold less BlpI-HiBiT than BlpAB (Fig. 5C). Therefore, RtgAB and ComAB/BlpAB do not efficiently cross-secrete each other's substrates. Consistent with this, RtgAB<sup>+</sup> strains do not show

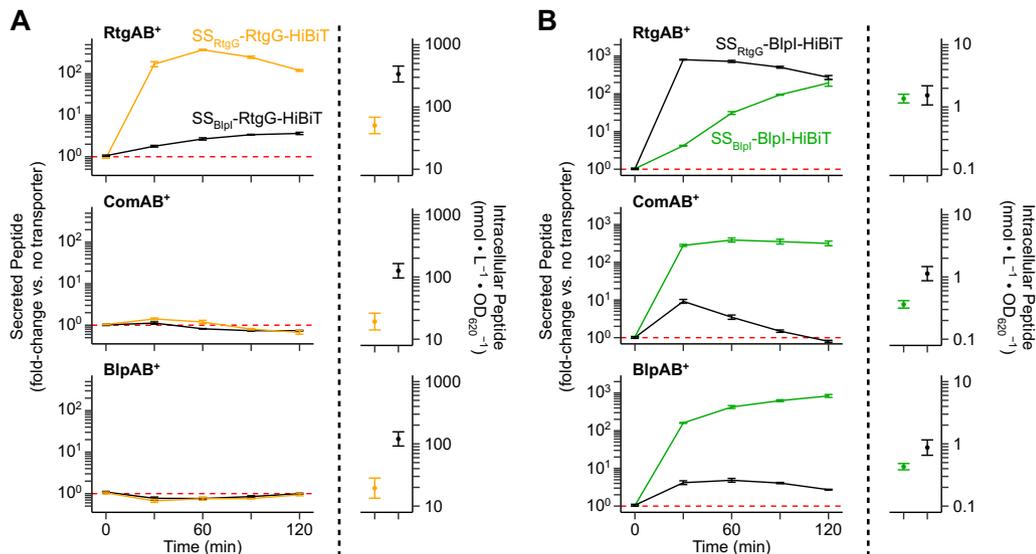


**FIG 5** RtgAB secretes different GG peptides from ComAB and BlpAB. R6 ComAB<sup>-</sup>/BlpAB<sup>-</sup>/RtgAB<sup>-</sup> and single-transporter-positive ComAB<sup>+</sup>, BlpAB<sup>+</sup>, and RtgAB<sup>+</sup> strains expressing RtgC-HiBiT (A), RtgG-HiBiT (B), or BlpI-HiBiT (C) were grown in CDM<sup>+</sup> to an OD<sub>620</sub> of 0.05 and treated with 200 ng/ml CSP, 200 ng/ml BlpC, and 20 nM RtgS<sub>A</sub>-C10. Samples were taken every 30 min, and extracellular HiBiT signal was quantified. Data are presented as fold change values relative to the ComAB<sup>-</sup>/BlpAB<sup>-</sup>/RtgAB<sup>-</sup> control. Red dashed lines represent fold change = 1 (no difference versus the control). Plots show means ± SEs from 3 independent experiments.

differences in *com* or *blp* activation compared to that of RtgAB<sup>-</sup> strains during growth in CDM<sup>+</sup> (see Fig. S5). 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.

**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 nonsubstrate GG peptides. We showed that the BlpI signal sequence ( $SS_{BlpI}$ ) prevents secretion of the RtgG cargo peptide through RtgAB (Fig. 6A). 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 ( $SS_{RtgG}$ ) promotes secretion of the BlpI cargo peptide through RtgAB while preventing its secretion through ComAB and BlpAB (Fig. 6B).

To rule out the possibility of differences in peptide expression being solely responsible for the secretion differences, we also measured the amount of intracellular peptide in each assay (Fig. 6A and B; right-hand graphs). The signal sequence swaps affected 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<sup>+</sup> versus ComAB<sup>+</sup> versus BlpAB<sup>+</sup>), secretion was not. Thus, the observed changes in secretion between the different peptides most likely reflect differences in peptide-transporter interactions.



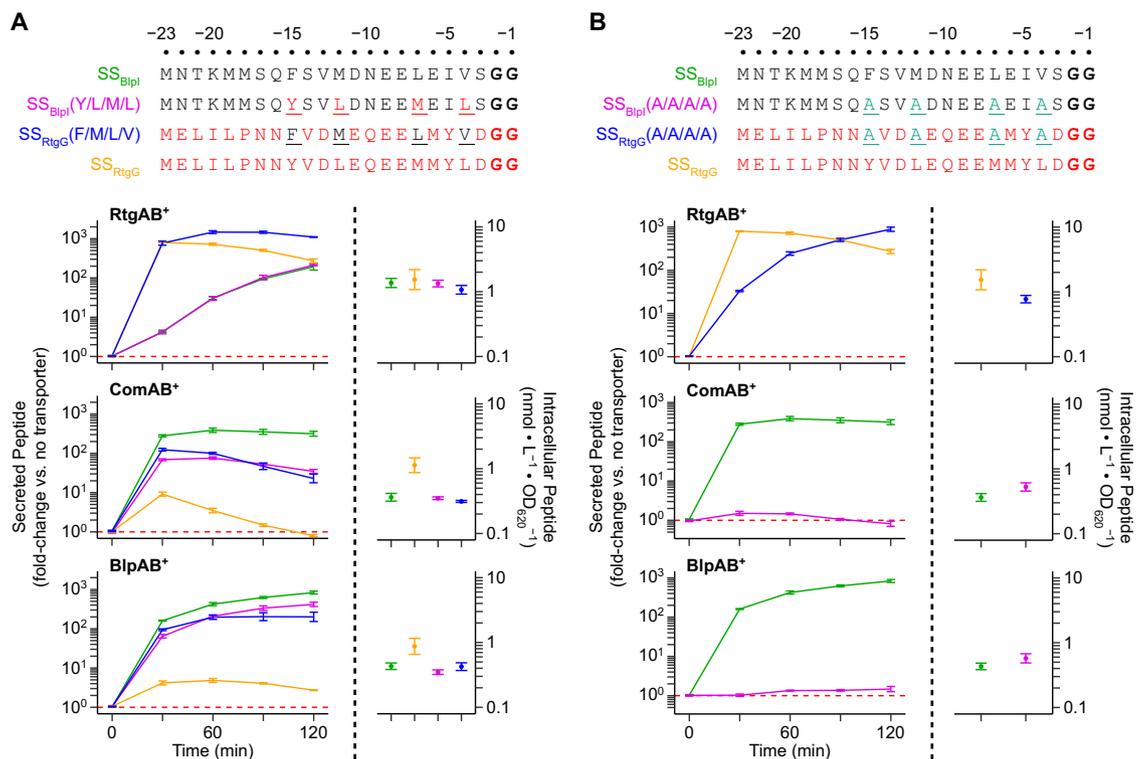
**FIG 6** RtgAB, ComAB, and BlpAB recognize GG peptides through their N-terminal signal sequences. R6 ComAB<sup>-</sup>/BlpAB<sup>-</sup>/RtgAB<sup>-</sup> and single-transporter-positive ComAB<sup>+</sup>, BlpAB<sup>+</sup>, and RtgAB<sup>+</sup> strains expressing signal sequence-swapped RtgG-HiBiT cargo peptide (A) or BlpI-HiBiT cargo peptide (B) were grown in CDM+ to an OD<sub>620</sub> of 0.05 and treated with 200 ng/ml CSP, 200 ng/ml BlpC, and 20 nM RtgS<sub>A</sub>-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<sup>-</sup>/BlpAB<sup>-</sup>/RtgAB<sup>-</sup> control. Red dashed lines represent fold change = 1 (no difference versus the control). At the 120-min time point, intracellular peptide content was also quantified (right). Plots show means ± SEs from 3 independent experiments.

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 own 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 were previously implicated in substrate recognition by PCATs (27, 28, 30). The combination of the four residues at these positions from SS<sub>BlpI</sub> (F/M/L/V) introduced into SS<sub>RtgG</sub> promote secretion through ComAB/BlpAB, although they were not strictly required for secretion in the context of SS<sub>BlpI</sub> (Fig. 7A). The complementary association did not hold for RtgAB-mediated secretion, in that the four residues from SS<sub>RtgG</sub> (Y/L/M/L) were neither necessary nor sufficient for secretion through RtgAB (Fig. 7A). Additionally, alanine substitutions at all four positions in SS<sub>RtgG</sub> only partially impeded secretion through RtgAB, but the same substitutions in SS<sub>BlpI</sub> prevented secretion through ComAB and BlpAB almost entirely (Fig. 7B).

#### **A specific motif at the N-terminal ends of *rtg* GG peptide signal sequences promotes secretion through RtgAB.**

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 in ComAB/BlpAB substrates. Residue swaps at positions -22 to -18 in SS<sub>RtgG</sub> and SS<sub>BlpI</sub> demonstrated that secretion through RtgAB, but not ComAB or BlpAB, depends on specific signal sequence residues in this region (Fig. 8A). The P(-18)M substitution in SS<sub>RtgG</sub> modestly decreased secretion through RtgAB, and removal of all residues on the N-terminal side of this substitution further decreased secretion (Fig. 8B). Meanwhile, removal of the residues at the same positions from SS<sub>BlpI</sub> did not change secretion through RtgAB (Fig. 8C). These data indicate that the residues in this region in SS<sub>RtgG</sub> were selected to interact with RtgAB rather than to avoid steric clash. Alanine scanning mutagenesis of the -22 to -19 region of SS<sub>RtgG</sub> revealed that secretion through RtgAB was not sensitive to mutation at any single site (see Fig. S6). These data can be explained by

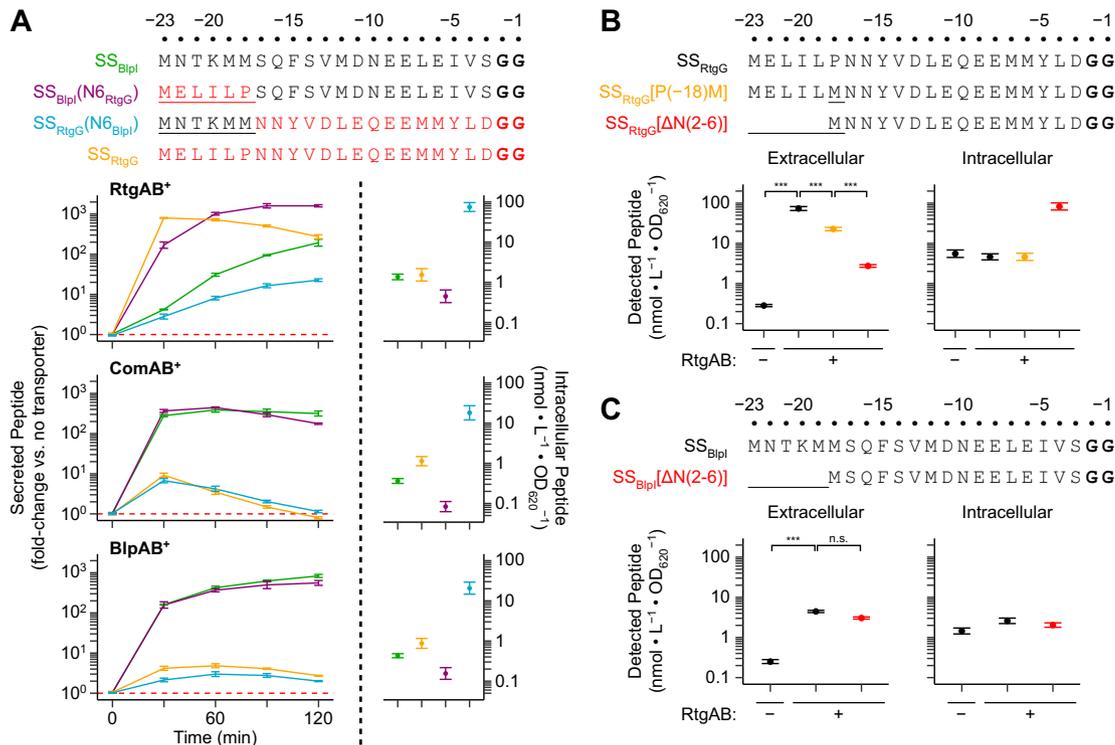


**FIG 7** Specific GG peptide signal sequence residues at positions  $-15$ ,  $-12$ ,  $-7$ , and  $-4$  modulate secretion by ComAB and BlpAB but not RtgAB. R6 ComAB<sup>-</sup>/BlpAB<sup>-</sup>/RtgAB<sup>-</sup> and single-transporter-positive ComAB<sup>+</sup>, BlpAB<sup>+</sup>, and RtgAB<sup>+</sup> strains expressing BlpI or RtgG signal sequences with residue swaps (A) or alanine substitutions (B) at positions  $-15$ ,  $-12$ ,  $-7$ , and  $-4$  (top, mutated positions are underlined) fused to BlpI-HiBiT cargo peptide were grown in CDM+ to an OD<sub>620</sub> of 0.05 and treated with 200 ng/ml CSP, 200 ng/ml BlpC, and 20 nM RtgS<sub>A</sub>-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<sup>-</sup>/BlpAB<sup>-</sup>/RtgAB<sup>-</sup> control. Red dashed lines represent fold change = 1 (no difference versus the control). At the 120-min time point, intracellular peptide content was also quantified (right). Plots show means  $\pm$  SEs from 3 independent experiments.

multiple redundant residues mediating the interactions in this region or the interactions being tolerant to alanine substitution. We conclude that RtgAB recognizes *rtg* GG peptides through interactions involving the signal sequence residues in the  $-22$  to  $-18$  region. At the same time, RtgAB's substrate recognition mechanism has evolved to be less reliant than that of ComAB or BlpAB on interactions with the hydrophobic signal sequence residues at positions  $-15$ ,  $-12$ ,  $-7$ , and  $-4$ .

**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 (12). Rgg/SHP and ComR/S circuits can act as either cell density-dependent quorum-sensing systems (12) or timing devices (40). Our data suggest RtgR/S behaves like the former (see Fig. S7 in the supplemental material). A purely intracellular signaling pathway has been reported for XIP in *Streptococcus mutans* (14, 41). Such a pathway is unlikely to exist for RtgR/S, since *rtg* autoinduction requires both PptAB and Ami (Fig. S1A). 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 (12). 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 (12, 20). RtgR is phylogenetically closer to the ComRs than SHP-associated Rgg regulators but does not cluster with either group (12). Using a published list of Rgg regulators (12), we found two RtgR-like regulators associated with Trp-X-Trp (WxW)



**FIG 8** A unique motif found at the N-terminal ends of the *rtg* GG peptides promotes secretion by RtgAB. (A) R6 ComAB<sup>-</sup>/BlpAB<sup>-</sup>/RtgAB<sup>-</sup> and single-transporter-positive ComAB<sup>+</sup>, BlpAB<sup>+</sup>, and RtgAB<sup>+</sup> strains expressing BlpI or RtgG signal sequences with residue swaps at positions -22 to -18 (top, mutated positions are underlined) fused to BlpI-HiBiT cargo peptide were grown in CDM+ to an OD<sub>620</sub> of 0.05 and treated with 200 ng/ml CSP, 200 ng/ml BlpC, and 20 nM RtgS<sub>A</sub>-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<sup>-</sup>/BlpAB<sup>-</sup>/RtgAB<sup>-</sup> control. Red dashed lines represent fold change = 1 (no difference versus the control). At the 120-min time point, intracellular peptide content was also quantified (right). Plots show means ± SEs from 3 independent experiments. R6 ComAB<sup>-</sup>/BlpAB<sup>-</sup>/RtgAB<sup>-</sup> and RtgAB<sup>+</sup> strains expressing mutated SS<sub>RtgG</sub> (B) or SS<sub>BlpI</sub> (C) (top, mutated positions are underlined) fused to BlpI-HiBiT cargo peptide were grown and treated with pheromone as in panel A. Extracellular (left) and intracellular (right) HiBiT signals were quantified at 60 min posttreatment. Plots show means ± SEs from 3 independent experiments. n.s., not significant; \*\*\*, *P* < 0.001 by ANOVA with Tukey's HSD.

motif-containing pheromones: SPD<sub>1518</sub> (Rgg1518) from *S. pneumoniae* D39 and SSA<sub>2251</sub> from *Streptococcus sanguinis* SK36 (predicted unprocessed pheromone sequences, MGFKKYLNLPKNSGFLIWSWIQLIWFETWFWG and MKKIVYNLILLAVTSIVTTSVFP WWWLWW, respectively). Expression analysis of the pheromone operon associated with *rgg1518* using PneumoExpress (42) revealed that the pheromone and genes SPD<sub>1513</sub> to SPD<sub>1517</sub> 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<sup>+</sup> 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 harbor a functional RtgR and therefore can respond to pheromone, only 17% of strains are RtgAB<sup>+</sup>. Most RtgAB<sup>-</sup> strains still harbor 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 that five of the seven *rtg* GG peptide genes are always associated with downstream genes encoding hypothetical proteins that resemble bacteriocin immunity proteins (43). The fact that the Sp9-BS68 strain with a functional *rtg* locus demonstrated a competitive advantage over

the  $\Delta rtgR \Delta rtgS1$  strain during dual infection is consistent with the bacteriocin hypothesis, as the regulator mutant would be unable to upregulate immunity, although we cannot exclude that other *rtgR*-regulated factors play a role in this fitness advantage. Regardless of the specific function of the *rtg* GG peptides, the fact that most RtgAB<sup>-</sup> strains are still RtgR<sup>+</sup> 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 nonfunctional RtgAB.

The case of RtgAB and ComAB/BlpAB allowed us to study how two sets of PCATs which coexist in the same strain preferentially secrete different sets of peptides through slight differences in substrate recognition. Unlike ComAB and BlpAB, RtgAB recognizes its substrates partially using a motif located at the N-terminal ends 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 (28, 30) or can be missing entirely (33, 44, 45). As far as we are aware, RtgAB is unique among PCATs in recognizing a signal sequence motif located so distantly from the cleavage site. Future efforts will be directed toward identifying the specific nature of the interaction between the N-terminal motif and RtgAB and the exact signal sequence residues involved.

The insights into the sequence determinants of PCAT substrate selectivity gained here illuminate 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 coregulated PCAT (18). In these cases, it would be helpful to have sequence-based approaches for assigning potential transporters to these “orphan” GG peptides. Moreover, for strains that harbor 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 (3), GG peptides have now been linked to biofilm formation, colonization of host niches, and dissemination during infection (18, 46). Ultimately, the ability to predict and rationalize PCAT-GG peptide pairings will advance our understanding of a broad range of biologically significant microbial processes.

## MATERIALS AND METHODS

**Strains and growth conditions.** All strains were derived from Sp9-BS68 (36), D39, or the R6 strain P654 (referred to as PSD100 in reference 47) (see Table S1 and methods in Text S1 in the supplemental material for details). The modified R6 strain was used for some *in vitro* assays because previous work demonstrating the *blp-com* connection was performed in this strain background. Pneumococcus was grown in either filter-sterilized THY (Todd Hewitt broth plus 0.5% yeast extract) or CDM+ (see methods in Text S1) (38) at 37°C. All media contained 5  $\mu$ g/ml catalase. All CDM+ was supplemented with 0.5% (vol/vol) THY. Except where noted otherwise, pneumococcal cultures used for experiments were inoculated to an OD<sub>620</sub> of 0.0015 from starter cultures grown in THY (pH 7.4) to an OD<sub>620</sub> of 0.275 and frozen at -80°C in 13% glycerol. Starter cultures were pelleted at 6,000  $\times g$  for 5 min at room temperature and resuspended in the appropriate growth medium for the experiment before being used for inoculation. Antibiotics were used at the following concentrations: chloramphenicol, 2  $\mu$ g/ml; gentamicin, 200  $\mu$ g/ml; kanamycin, 500  $\mu$ g/ml; spectinomycin, 200  $\mu$ g/ml; streptomycin, 100  $\mu$ g/ml.

**Transformations.** Transformation protocols were adapted from those described in reference 48. See methods in Text S1 and Table S1 for details and primers used for constructing transforming DNA products. Unmarked chromosomal mutations were created via Janus (49), Sweet Janus (50), or Janus2 (Text S1) exchange. Transformants were verified by Sanger sequencing.

**Luciferase reporter time course assays.** For *com-blp* activation assays only, starter cultures were grown in THY (pH 6.8) to an OD<sub>620</sub> of 0.075 to prevent *com-blp* activation. Cells were grown in THY or CDM+ in a white, clear-bottom 96-well plate (655098; Greiner Bio-One), 200  $\mu$ l per well. For assays using

firefly luciferase, the following concentrations of firefly luciferin (88294; Thermo Fisher Scientific) were added to the medium: 330  $\mu\text{M}$  (single reporter and dual reporter, CDM+), 165  $\mu\text{M}$  (dual reporter, THY). For assays using NanoLuc luciferase, the following concentrations of Nano-Glo substrate (N1121; Promega) were added to the media: 1:5,000 (CDM+), 1:10,000 (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 reference 25. For locus activation assays, timings of activation events were calculated as described in the methods in Text S1 and compared using survival analysis. 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.

**RtgS dose-response assays.** Cells expressing  $P_{rtgS1}$ -*luc* reporters were grown in THY or CDM+ containing 330  $\mu\text{M}$  firefly luciferin. At an  $\text{OD}_{620}$  of 0.02, cultures were aliquoted into a white, clear-bottom 96-well plate (655098; Greiner Bio-One), 100  $\mu\text{l}$  per well. Each well of the plate was pre-filled with 100  $\mu\text{l}$  sterile medium containing 0.5% (vol/vol) dimethyl sulfoxide (DMSO), 330  $\mu\text{M}$  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.

**Peptide secretion assays.** Cells were inoculated from starter cultures to an  $\text{OD}_{620}$  of 0.005 and grown in CDM+. At an  $\text{OD}_{620}$  of 0.05, cells were treated with 200 ng/ml CSP1, 200 ng/ml  $\text{BlpC}_{\text{R6}}$ , and 20 nM  $\text{RtgS}_{\Delta\text{C10}}$ . Samples were taken for HiBiT quantification at appropriate time points. For native  $\text{BlpI}$ -HiBiT assays only, clarified supernatants were obtained after centrifugation at  $6,000 \times g$  for 5 min at 4°C. For all other assays, cells were retained in the samples. HiBiT signal was quantified by mixing samples with HiBiT extracellular detection reagent (N2421; Promega) at a 1:1 ratio and reading luminescence with a Synergy HTX plate reader. 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. Extracellular peptide was removed from these samples by proteinase K digestion, and then the cells were lysed and HiBiT signal was quantified as described above. Standards consisting of synthetic L10-HiBiT peptide (25) mixed with samples of a non-HiBiT-expressing strain were used to generate standard curves to use for calculating HiBiT-tagged peptide concentrations in experimental samples. See methods in Text S1 for more details. Differences between groups were assessed by analysis of variance (ANOVA) using the emmeans package (v1.2.3) in R.

**Genomic analysis of *rtg*.** Analysis of *rtg* was performed using the assembled genomes of the Massachusetts isolate collection (BioProject accession [PRJEB2632](#)). See methods in Text S1 for details.

**Mouse colonization assays.** Mouse colonization was performed as described in reference 25. Briefly, dual or single-strain mixtures of Sp9-BS68 were inoculated into the nasopharynx of unanaesthetized 5- to 7-week-old female BALB/c mice (Taconic) with  $1.0 \times 10^6$  to  $3.0 \times 10^6$  CFU/mouse in 10  $\mu\text{l}$  of sterile phosphate-buffered saline (PBS). For dual inoculated mice, the ratio of the kanamycin-resistant strain to the spectinomycin-resistant strain was between 0.25 and 0.6. Mice were euthanized with  $\text{CO}_2$  overdose after 72 h, and nasopharyngeal colonization was sampled by nasal wash. See methods in Text S1 for IACUC approval and details on how colonization density and competitive indices were calculated. 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.

**Data availability.** Sequences of Janus+ and Janus2 constructs were deposited in GenBank under accession numbers [MN848328](#) and [MN848329](#), respectively. The *rtg* locus from Sp9-BS68 including the new sequencing that allowed us to connect existing contigs and *rtg* gene designations established here was deposited as accession number [MN848330](#).

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**TEXT S1**, PDF file, 0.3 MB.

**FIG S1**, PDF file, 0.6 MB.

**FIG S2**, PDF file, 0.5 MB.

**FIG S3**, PDF file, 0.8 MB.

**FIG S4**, PDF file, 0.4 MB.

**FIG S5**, PDF file, 0.5 MB.

**FIG S6**, PDF file, 0.4 MB.

**FIG S7**, PDF file, 0.5 MB.

**TABLE S1**, PDF file, 0.2 MB.

## ACKNOWLEDGMENTS

This work was supported by funding from the National Institutes of Health T32 AI007528 (to C.Y.W.), R01 AI101285 (to S.D.), and R56 AI101285 (to S.D.).

## REFERENCES

- Economou A. 2002. Bacterial secretome: the assembly manual and operating instructions (review). *Mol Membr Biol* 19:159–169. <https://doi.org/10.1080/09687680210152609>.
- Green ER, Mecsas J. 2016. Bacterial secretion systems: an overview. *Microbiol Spectr* 4:VMBF-0012-2015. <https://doi.org/10.1128/microbiolspec.VMBF-0012-2015>.
- 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* 16:229–240. <https://doi.org/10.1111/j.1365-2958.1995.tb02295.x>.
- Gebhard S. 2012. ABC transporters of antimicrobial peptides in *Firmicutes* bacteria - phylogeny, function and regulation. *Mol Microbiol* 86:1295–1317. <https://doi.org/10.1111/mmi.12078>.
- Chikindas ML, Weeks R, Drider D, Chistyakov VA, Dicks LM. 2018. Functions and emerging applications of bacteriocins. *Curr Opin Biotechnol* 49:23–28. <https://doi.org/10.1016/j.copbio.2017.07.011>.
- Hui FM, Zhou L, Morrison DA. 1995. Competence for genetic transformation in *Streptococcus pneumoniae*: organization of a regulatory locus with homology to two lactococcal A secretion genes. *Gene* 153:25–31. [https://doi.org/10.1016/0378-1119\(94\)00841-f](https://doi.org/10.1016/0378-1119(94)00841-f).
- 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:e00179-11. <https://doi.org/10.1128/mBio.00179-11>.
- Mendonça 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. <https://doi.org/10.3389/fmicb.2016.02156>.
- Petersen FC, Scheie AA. 2000. Genetic transformation in *Streptococcus mutans* requires a peptide secretion-like apparatus. *Oral Microbiol Immunol* 15:329–334. <https://doi.org/10.1034/j.1399-302x.2000.150511.x>.
- 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* 63:1209–1222. <https://doi.org/10.1111/j.1365-2958.2007.05581.x>.
- 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* 65:652–658. <https://doi.org/10.1128/AEM.65.2.652-658.1999>.
- 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* 80:1102–1119. <https://doi.org/10.1111/j.1365-2958.2011.07633.x>.
- 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:e02306-14. <https://doi.org/10.1128/mBio.02306-14>.
- Chang JC, Federle MJ. 2016. PptAB exports Rgg quorum-sensing peptides in *Streptococcus*. *PLoS One* 11:e0168461. <https://doi.org/10.1371/journal.pone.0168461>.
- 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* 188:7230–7241. <https://doi.org/10.1128/JB.00877-06>.
- 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. <https://doi.org/10.1038/s41598-018-24910-1>.
- Junges R, Salvadori G, Shekhar S, Åmdal HA, Periselneris JN, Chen T, Brown JS, Petersen FC, 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:e00324-17. <https://doi.org/10.1128/mSphere.00324-17>.
- 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. <https://doi.org/10.1111/mmi.13721>.
- 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* 192:1444–1454. <https://doi.org/10.1128/JB.01251-09>.
- 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. <https://doi.org/10.1111/j.1365-2958.2010.07361.x>.
- 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* 173:372–381. <https://doi.org/10.1128/jb.173.1.372-381.1991>.
- 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* 182:4696–4703. <https://doi.org/10.1128/jb.182.17.4696-4703.2000>.
- 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. <https://doi.org/10.1371/journal.ppat.1005422>.
- 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. <https://doi.org/10.1371/journal.ppat.1005413>.
- 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. <https://doi.org/10.1073/pnas.1804668115>.
- 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. <https://doi.org/10.3389/fcimb.2019.00055>.
- 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* 47:2531–2538. <https://doi.org/10.1021/bi702253n>.
- Bobeica SC, Dong S-H, 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:e42305. <https://doi.org/10.7554/eLife.42305>.
- 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 pre-peptide. *Peptides* 30:1412–1420. <https://doi.org/10.1016/j.peptides.2009.05.021>.
- 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* 285:10777–10785. <https://doi.org/10.1074/jbc.M109.093781>.
- Lin DY, Huang S, Chen J. 2015. Crystal structures of a polypeptide processing and secretion transporter. *Nature* 523:425–430. <https://doi.org/10.1038/nature14623>.
- 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* 80:6647–6655. <https://doi.org/10.1128/AEM.02312-14>.
- 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* 126:23–29. <https://doi.org/10.1016/j.jbiosc.2018.01.015>.
- 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* 161:2069–2078. <https://doi.org/10.1099/mic.0.000157>.
35. Furgerson Ihnken LA, Chatterjee C, van der Donk WA. 2008. *In vitro* reconstitution and substrate specificity of a lantibiotic protease. *Biochemistry* 47:7352–7363. <https://doi.org/10.1021/bi800278n>.
  36. 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* 189:8186–8195. <https://doi.org/10.1128/JB.00690-07>.
  37. 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* 209:1781–1791. <https://doi.org/10.1093/infdis/jit818>.
  38. Marks LR, Reddinger RM, Hakansson AP. 2012. High levels of genetic recombination during nasopharyngeal carriage and biofilm formation in *Streptococcus pneumoniae*. *mBio* 3:e00200-12. <https://doi.org/10.1128/mBio.00200-12>.
  39. 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* 45:656–663. <https://doi.org/10.1038/ng.2625>.
  40. 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* 195:1845–1855. <https://doi.org/10.1128/JB.02196-12>.
  41. 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:e00444-18. <https://doi.org/10.1128/mSphere.00444-18>.
  42. 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. <https://doi.org/10.1093/nar/gky750>.
  43. Nes IF, Diep DB, Håvarstein LS, Brurberg MB, Eijsink V, Holo H. 1996. Biosynthesis of bacteriocins in lactic acid bacteria. *Antonie Van Leeuwenhoek* 70:113–128. <https://doi.org/10.1007/bf00395929>.
  44. 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* 76:483–492. <https://doi.org/10.1128/AEM.01597-09>.
  45. 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* 72:3383–3389. <https://doi.org/10.1128/AEM.72.5.3383-3389.2006>.
  46. 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. <https://doi.org/10.1371/journal.ppat.1007328>.
  47. 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* 195:1561–1572. <https://doi.org/10.1128/JB.01964-12>.
  48. Dawid S, Sebert ME, Weiser JN. 2009. Bacteriocin activity of *Streptococcus pneumoniae* is controlled by the serine protease HtrA via posttranscriptional regulation. *J Bacteriol* 191:1509–1518. <https://doi.org/10.1128/JB.01213-08>.
  49. 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* 67:5190–5196. <https://doi.org/10.1128/AEM.67.11.5190-5196.2001>.
  50. 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. <https://doi.org/10.1371/journal.pone.0100510>.