The molecular mechanism of N-acetylglucosamine side-chain attachment to the Lancefield group A Carbohydrate in *Streptococcus pyogenes*.

Jeffrey S. Rush[a], Rebecca J. Edgar[a], Pan Deng[b], Jing Chen[a], Haining Zhu[a], Nina M. van Sorge[c], Andrew J. Morris[b], Konstantin V. Korotkov[a], and Natalia Korotkova[a#]

Department of Molecular and Cellular Biochemistry, University of Kentucky, Lexington, Kentucky, USA[a]; Division of Cardiovascular Medicine and the Gill Heart Institute, University of Kentucky Lexington, Kentucky, USA[b]; Department of Medical Microbiology, University Medical Center Utrecht, Utrecht, The Netherlands[c].

[#]Address correspondence to Natalia Korotkova, nkorotkova@uky.edu

Running Head: Mechanism of cell wall modification with GlcNAc in GAS

The word count for the abstract: 196

The word count for the text: 12387

Abstract

Streptococcus pyogenes or Group A *Streptococcus* (GAS) causes numerous diseases in humans, ranging from minor skin and throat infections, to life-threatening systemic conditions such as bacteremia, necrotizing fasciitis, and toxic-shock syndrome. GAS synthesizes a key antigenic surface polymer — the Lancefield group A Carbohydrate (GAC). GAC is attached to peptidoglycan and consists of a polyrhamnose polymer, with N-acetylglucosamine (GlcNAc) side chains, that is an essential virulence determinant. The molecular details of the mechanism of polyrhamnose modification with GlcNAc are currently unknown. In this report we demonstrate that GAC biosynthesis requires two distinct undecaprenol-linked GlcNAc-lipid intermediates: GlcNAc-pyrophosphoryl-undecaprenol (GlcNAc-P-P-Und) produced by GacO and GlcNAc-phosphate-undecaprenol (GlcNAc-P-Und) produced by GacI. The GAC polyrhamnose backbone is assembled on GlcNAc-P-P-Und. Furthermore, our data suggests that GlcNAc-P-Und is used by the membrane glycosyltransferase GacL to transfer GlcNAc from GlcNAc-P-Und to the polyrhamnose polysaccharide. In addition, GacJ, a small membrane protein, forms a complex with GacI and significantly stimulates its catalytic activity. We show that GlcNAc modification of polyrhamnose protects GAS from amidase-induced lysis. Thus, our study significantly expands our understanding of the biosynthesis of GAS cell wall polysaccharide and points to the functional importance of polysaccharide modifications in protection of peptidoglycan from lytic enzymes.

IMPORTANCE

In many species of Lactobacillales peptidoglycan is decorated by polyrhamnose polysaccharides that are critical for cell envelope integrity and cell shape and represent key antigenic determinants. Despite the biological importance of these polysaccharides, their biosynthetic pathways have received limited study. The important human pathogen, *Streptococcus pyogenes*, synthesizes a cell wall-associated polyrhamnose polysaccharide with N-acetylglucosamine (GlcNAc) side chains. The GlcNAc side chain is an important virulence determinant used by bacteria to evade the host innate immune defense. Here we identify the molecular mechanism of polyrhamnose modification with GlcNAc in *S. pyogenes*. We show that GacI synthesizes GlcNAc-phosphate-undecaprenol (GlcNAc-P-Und) aided by GacJ. GlcNAc-P-Und is apparently used by

GacL to transfer GlcNAc to polyrhamnose. We demonstrate that GacI homologs perform a similar function in *Streptococcus agalactiae* and *Enterococcus faecalis*. Thus, the elucidation of polysaccharide biosynthesis in *S. pyogenes* enhances our understanding of how other Gram-positive bacteria produce essential components of their cell wall.

INTRODUCTION

The cytoplasmic membrane of Gram-positive bacteria is surrounded by a thick cell wall consisting of multiple peptidoglycan layers decorated with proteins and a variety of carbohydrate-based polymers. Rhamnose (Rha) is the main component of the carbohydrate structures in many species of the Lactobacillales order (1). These polymers play essential roles in maintaining and protecting bacterial cell envelopes and in pathogenesis of infections caused by *Streptococcus pyogenes* and *Enterococcus faecalis* (1). It is of paramount importance to understand the molecular mechanisms of Rha-containing cell wall polysaccharide biosynthesis for developing novel therapeutics against these bacterial pathogens.

S. pyogenes or Group A *Streptococcus* (GAS) is associated with numerous diseases in humans ranging from minor skin and throat infections such as impetigo and pharyngitis to life-threatening invasive infections such as streptococcal toxic syndrome and necrotizing fasciitis (2). The main component of GAS cell wall is the Lancefield group A Carbohydrate (GAC) that comprises about 40-60% of the total cell wall mass (3). Serological grouping of beta-hemolytic streptococci (A, B, C, E, F, and G groups), introduced by Rebecca Lancefield in 1933, is based on the detection of carbohydrate antigens present on the cell wall (4). GAC is presumably covalently linked to N-acetylmuramic acid (MurNAc) of peptidoglycan, although the molecular structure of the linkage unit has not been established (5). All serotypes of GAS produce GAC consisting of a polyrhamnose backbone with N-acetylglucosamine (GlcNAc) side chains (6, 7). The average molecular mass of GAC has been reported to be 8.9±1.0 kDa, corresponding to an average of 18 repeating units of [→3)α-Rha(1→2)[β-GlcNAc(1→3)]α-Rha(1→] (8). Since Rha is absent in mammalian cells, GAC is an attractive candidate for a universal GAS vaccine (8, 9). Moreover, the GlcNAc side chains are an important virulence

determinant in GAS (5). GAS mutants lacking GlcNAc are susceptible to innate immune clearance by neutrophils and antimicrobial agents, and are significantly attenuated in animal models of GAS infection (5).

Similar to the biosynthesis of other cell-envelope polymers — peptidoglycan, lipopolysaccharide, wall teichoic acid (WTA), and capsular polysaccharide — the synthesis of Rha-containing carbohydrate structures is likely initiated on the inside of the cytoplasmic membrane and proceeds through several steps including the attachment of the first sugar residue to a lipid carrier, undecaprenyl phosphate (Und-P), followed by elongation of the polysaccharide through stepwise addition of activated sugar residues to the lipid carrier (1). After translocation of the polysaccharide across the membrane and further modifications, it is presumably attached to peptidoglycan via a phosphate ester linkage (1). The first membrane step of O-antigen biosynthesis in enterobacteria and WTA biosynthesis in *Bacillus subtilis* and *Staphylococcus aureus* (10) is catalyzed by the UDP-GlcNAc:Und-P GlcNAc-1-phosphate transferase encoded by WecA homologs (11, 12). WecA transfers GlcNAc-phosphate from UDP-GlcNAc to Und-P, forming GlcNAc-pyrophosphoryl-undecaprenol (GlcNAc-P-P-Und). Although GAS strains do not produce WTA, all GAS genomes contain a *wecA* homolog, *gacO* (5). Significantly, GAC biosynthesis is sensitive to tunicamycin, a known inhibitor of WecA (5) and the *Streptococcus mutans* GacO homolog, RgpG, has been shown to complement WecA activity in *E. coli* (13). These observations support an essential role for GacO in the biosynthesis of GlcNAc-P-P-Und and suggest GlcNAc-P-P-Und may function as a lipid-anchor in the initiation of GAC biosynthesis.

Other genes required for GAC biosynthesis and transport are found in a separate location on the GAS chromosome and comprise a 12-gene locus (*gacA*-*gacL*) (Fig.1A). The first three genes, *gacA*-*gacC* together with *gacG,* are conserved in many species of the Lactobacillales order (1). *GacA* encodes a dTDP-4-dehydrorhamnose reductase, the enzyme responsible for dTDP-rhamnose biosynthesis (14). *GacB*, *gacC*, and *gacG* encode putative cytoplasmic rhamnosyltransferases. In *S. mutans* homologous genes *rgpA*, *rgpB* and *rgpF* are involved in rhamnan backbone biosynthesis (15). In GAS, *gacA*-*gacC* are essential for viability and cannot be deleted (5). However, deletion mutants have been obtained in other genes of the GAC gene cluster (5). It has been

shown that *gacI*, *gacJ* and *gacK* are non-essential for viability, but are required for GlcNAc side-chain addition to the polyrhamnose (5). They encode a putative cytoplasmic glycosyltransferase, a small membrane protein and a protein with homology to the Wzx family of membrane proteins involved in the export of O-antigen and teichoic acids, respectively. In contrast, inactivation of *gacD*, *gacE*, *gacF*, *gacG*, *gacH* or *gacL* has no effect on GAS viability and the GAC produced by these mutants is reported to display a wild type (WT) antigenic profile, indicating the presence of the immunodominant GlcNAc side chains (5). *GacD* and *gacE* encode the components of an ABC transport system. In *S. mutans* the *gacD* and *gacE* homologs are responsible for rhamnan polysaccharide transport (15). *GacF*, *gacH* and *gacL* encode a cytosolic glycosyltransferase, a putative membrane-associated glycerol phosphate transferase and a membrane-associated glycosyltransferase, respectively.

In this study we address the molecular mechanism of GlcNAc attachment to polyrhamnose. We demonstrate that GacI catalyzes formation of GlcNAc-P-Und, and GacJ stimulates the catalytic activity of GacI. Subsequently GacL transfers GlcNAc from GlcNAc-P-Und to the polyrhamnose backbone of GAS polysaccharide. Moreover, we confirm GacO function in GlcNAc-P-P-Und formation and demonstrate the role of this GlcNAc-lipid in initiation of polyrhamnose biosynthesis.

RESULTS

GacL is required for GlcNAc attachment to polyrhamnose. GacL, a polytopic (12–13 transmembrane segments) membrane protein, is reported to be dispensable for GlcNAc attachment to polyrhamnose (5). To investigate the function of GacL in GAC biogenesis, we disrupted *gacL* in the hyperinvasive *S. pyogenes* M1T1 serotype strain, MGAS5005 (16), creating the 5005Δ*gacL* mutant. In agreement with published data (5) 5005Δ*gacL* did not display any detectable growth phenotype in comparison to the WT strain. Surprisingly, we found that 5005Δ*gacL* cells failed to bind GlcNAc-specific anti-GAC antibodies, suggesting a loss of the GlcNAc antigenic epitope (Fig. 1B). The 5005Δ*gacL* phenotype was restored by expressing the WT copy of *gacL* on the mutant chromosome (Fig. 1B). To confirm that 5005Δ*gacL* cells are deficient in GlcNAc addition to polysaccharide, we measured the binding of fluorescently labeled succinylated wheat

germ agglutinin (sWGA), a lectin that specifically binds non-reducing terminal β-GlcNAc residues, to WT and 5005Δ*gacL* cells. Deletion of *gacL* led to a significant decrease in binding of WGA to the 5005Δ*gacL* strain as compared with the WT (Fig. 1C), indicating that the mutant has substantially less GlcNAc-containing saccharides on the cell surface. Furthermore, direct compositional analysis of cell wall purified from WT and 5005Δ*gacL* cells confirmed a significant decrease of GlcNAc content in the 5005Δ*gacL* cell wall sample (Fig.1D, E and F). These data strongly support a role for GacL in GlcNAc side-chain attachment to polyrhamnose.

A *gacL* deletion mutant accumulates GlcNAc-P-Und. To investigate further the effect of GacL deletion in *S. pyogenes*, we isolated the phospholipid fractions from the membranes of WT and 5005Δ*gacL* cells. TLC analysis revealed accumulation of a previously unidentified phosphoglycolipid in 5005Δ*gacL* (Fig. 2A). The novel lipid was found to be stable to mild alkaline methanolysis, but sensitive to mild acid (0.1 N HCl, 50 ºC, 50% isopropanol) and reacted with orcinol spray (17), Dittmer-Lester phospholipid spray reagent (18) and with anisaldehyde, an isoprenol-specific reagent (19), consistent with its tentative identification as a glycophosphoprenol (data not shown). High-resolution negative ion ESI-MS analysis of the purified lipid identified a molecular ion $[M-H]^{-1}$ of M/z = 1048.73621 (Fig. 2B). In addition, prominent high-resolution fragment ions characteristic of PO_3, H_2PO_4, N-acetyl hexosamine-phosphate and Und-P were also found (Table 1). These data are consistent with a glycolipid comprised of N-acetyl hexosamine-phosphate-undecaprenol. In experiments described in detail below, we found that when membrane fractions from GAS are incubated with UDP-[^{3}H]GlcNAc, a [^{3}H]GlcNAc-lipid which co-chromatographs on TLC with this novel lipid is rapidly formed. Importantly, the product ion spectra shown in Figure 2B does not contain a fragment ion containing $[Und\text{-}PO_4\text{-}C_2H_2NHCOCH_3]^-$ (M/z=929.7), arising from a cross-ring fragmentation reaction, suggesting that the hydroxyl of the anomeric carbon and the nitrogen at the 2-position of the glycosyl ring are *trans* to the plane of the ring (20, 21). This observation strongly suggests that the anomeric hydroxyl is present in the β configuration. Furthermore, isolated [^{3}H]GlcNAc-P-Und phospholipid (described in detail below) is extremely sensitive to incubation with 50% phenol at 68 °C — a property that is consistent with β-GlcNAc-P-Und (22) (Table S4). Taken together, these data

identify the novel glycolipid accumulating in the 5005Δ*gacL* strain as β-GlcNAc-P-Und and suggest that GacL is a GlcNAc transferase using β-GlcNAc-P-Und as GlcNAc donor for the addition of the GlcNAc side-chains to GAC.

GAS synthesizes two GlcNAc-lipids *in vitro*. To investigate the biosynthetic origin of the novel glycolipid accumulating in 5005Δ*gacL* cells, membrane fractions from MGAS5005 were incubated with UDP-[^{3}H]GlcNAc and analyzed for lipid products as described in Material and Methods. Experiments showed that [^{3}H]GlcNAc was efficiently transferred from UDP-[^{3}H]GlcNAc into two detectable GlcNAc-lipids (Fig. 3A, Table 2). The major lipid product co-migrates on TLC with the novel lipid accumulating in the 5005Δ*gacL* strain and is identified as β-[^{3}H]GlcNAc-P-Und, as described above. The minor product is assumed to be [^{3}H]GlcNAc-P-P-Und, since it co-migrates with authentic GlcNAc-P-P-Und synthesized in *B. cereus* membranes (Fig. 3B) (23) and its formation is potently inhibited by tunicamycin (Fig. S1).

GacI encodes a UDP-GlcNAc:Und-P GlcNAc transferase activity. The GAS polysaccharide gene cluster contains a gene*, gacI*, reported to be required for GlcNAc addition to cell wall polysaccharide (5) and annotated as a GT-A type glycosyltransferase. The following studies were conducted to determine if GAS GacI might be responsible for the synthesis of GlcNAc-P-Und. To determine if GacI function was essential for formation of GlcNAc-P-Und, we generated a deletion of *gacI* in the MGAS5005. In agreement with published data (5) 5005Δ*gacI* lost reactivity with anti-GAC antibodies, indicating a loss of GlcNAc modification in polyrhamnose (Fig. 1B). When 5005Δ*gacI* membrane fractions were incubated with [^{3}H]UDP-GlcNAc no incorporation into [^{3}H]GlcNAc-P-Und was observed and only the minor lipid, [^{3}H]GlcNAc-P-P-Und was found (Fig. 3C, Table 2), strongly supporting an essential role for GacI in GlcNAc-P-Und synthesis. The observation that the 5005Δ*gacL* strain synthesizes normal levels of [^{3}H]GlcNAc-P-Und *in vitro* (Fig. 3D) indicates that 5005Δ*gacL* does not fail to add GlcNAc side-chains to polyrhamnose due to a lack of GlcNA-P-Und and supports the conclusion that GacL may be the GlcNAc-P-Und:polyRhamnan GlcNAc transferase.

To confirm the role of GacI in GlcNAc-P-Und biosynthesis we expressed GAS GacI in *E. coli*. When membrane fractions from *E. coli* carrying an empty vector were

incubated with UDP-[^{3}H]GlcNAc and analyzed by silica gel TLC, a small peak of [^{3}H]GlcNAc-P-P-Und was found (Table 2, Fig. S4A). In contrast, the membranes of recombinant *E. coli* expressing GacI accumulated two products corresponding to a small amount of GlcNAc-P-P-Und and a very large amount of GlcNAc-P-Und (Table 2, Fig. S4B). Thus, our data strongly indicate that GacI is the GlcNAc-P-Und synthase.

GacI homologs in GBS and *E. faecalis* function in the biosynthesis of GlcNAc-P-Und. The presence of GlcNAc-P-P-Und and GlcNAc-P-Und have been previously reported in *B. cereus* and *Bacillus megaterium* (12, 23). Our *in vitro* studies show that the MGAS5005 strain synthesizes two [^{3}H]GlcNAc-lipids that exactly co-migrate on TLC with the previously reported [^{3}H]GlcNAc-lipids from *B. cereus* (Fig. 3A and 3B). Moreover, an analysis of bacterial genomes using GacI in a BLAST search identified a GacI homolog (60% identity) in *B. cereus* and *B. megaterium*, suggesting its role in the GlcNAc-P-Und biosynthesis in these bacteria. Furthermore, *gacI* homologs were identified in the polysaccharide biosynthesis gene clusters of the important human pathogens *E. faecalis* (*epaI* with 48% sequence identity) (24) and *S. agalactiae* (Group B Streptococcus or GBS) (SAN_1536 gene with 46% sequence identity) (Fig. S2). To investigate whether the GBS homolog of GacI also catalyzes the formation of GlcNAc-P-Und, we engineered a knock out of the GacI homolog in GBS COH1. Although GBS COH1 membranes synthesize both [^{3}H]GlcNAc-P-P-Und and [^{3}H]GlcNAc-P-Und (Fig. 4B and Fig. S3A), GBS COH1Δ*gacI* membranes no longer synthesize [^{3}H]GlcNAc-P-Und (Table 2 and Fig. S3B).

To confirm that the GacI homolog detected in *E. faecalis*, EpaI (24), also possesses GlcNAc-P-Und synthase activity, EpaI was expressed exogenously in *E. coli* and found to actively catalyze the formation of [^{3}H]GlcNAc-P-Und (Table 2, Fig. S4C). Altogether, our results are consistent with the function of GacI homologs from GAS, GBS and *E. faecalis* in the transfer of GlcNAc from UDP-GlcNAc to Und-P forming GlcNAc-P-Und.

ATP stimulates GlcNAc-P-Und biosynthesis. Preliminary enzymatic properties for the synthesis of [^{3}H]GlcNAc-P-Und established that Mg^{2+} was the preferred divalent cation and that the formation of GlcNAc-P-Und was substantially stimulated by exogenously added Und-P (as a dispersion in CHAPS detergent). In early studies of GAC biosynthesis in GAS (25), it was shown that the addition of ATP dramatically

stimulated incorporation of radioactive UDP-[^{3}H]GlcNAc into [^{3}H]GlcNAc membrane lipids and [^{3}H]GlcNAc-polysaccharide. We confirmed that inclusion of 1 mM ATP significantly stimulated the incorporation of [^{3}H]GlcNAc into [^{3}H]GlcNAc-P-Und (Fig. 5 and Fig. S5), as well as [^{3}H]GlcNAc-P-P-Und (Fig. S5). However, the inclusion of ATP did not stimulate GacI activity in *in vitro* assays of CHAPS-soluble, affinity purified GacI solely dependent on exogenously added Und-P as acceptor. These data suggest that the effect of ATP addition is most likely due to formation of Und-P, *in situ,* by phosphorylation of endogenous undecaprenol by undecaprenol kinase (26).

GacJ forms a complex with GacI and enhances its catalytic efficiency. *GacJ* is located immediately downstream of *gacI* in the GAS GAC biosynthesis gene cluster and encodes a small 113 aa membrane protein (Fig. 1A). We hypothesized that GacI might form an obligate complex with GacJ. This hypothesis is based on the observation that *gacJ* and *gacI* are frequently located adjacent to each other on bacterial chromosomes and are sometimes fused to form a single polypeptide (accession numbers: GAM11018, ADH85075, ALC15489 and ADU67183). To test this hypothesis, we solubilized membranes from *E. coli* co-expressing GacJ and amino-terminal His-tagged GacI with the zwitterionic detergent CHAPS, and isolated GacI complexes using Ni-NTA chromatography (Fig. 6A). SDS-PAGE of the affinity-purified sample revealed two bands corresponding to the anticipated molecular sizes of GacI and GacJ (Fig. 6B). Proteomics analysis of the excised protein bands confirmed the identities of the recovered proteins. This result indicates that GacI and GacJ form a stable, CHAPS-soluble complex and co-purify during affinity chromatography.

To investigate if GacJ performs any catalytic function in association with GacI, we tested the proteins for GlcNAc-P-Und synthase activity *in vitro*, individually and in combination. When GacI was expressed singly in *E. coli*, enzymatically active protein was found in the membrane fraction (Table 2, Fig. S4B), indicating that GacI does not require GacJ for activity or membrane association. GacJ was found to be catalytically inactive when expressed by itself in *E. coli*, consistent with the observation that 5005Δ*gacI* membranes show no residual GlcNAc-P-Und synthase activity. Kinetic analysis of GlcNAc-P-Und synthase activity for Und-P in membrane fractions of *E. coli* expressing GacI revealed an apparent Km of 18.7 μM (Table 3, Fig. S6A). However, co-

expression of GacJ dramatically lowered the apparent Km of GacI for Und-P to 1.1 μM, suggesting a significant change in affinity for the lipid acceptor. Moreover, the V_{max} of GacI activity increased from 54.2 pmol/min/mg to 18.5 nmol/min/mg in the presence of GacJ (Table 3, Fig. S6B). A kinetic analysis of GlcNAc-P-Und synthase for Und-P in MGAS5005 strain gave an apparent Km of 6.4 μM and a V_{max} of 333 pmol/min/mg (Table 3, Fig. S6C), similar to the enzymatic parameters of GacI co-expressed with GacJ in *E. coli*. These enzymatic parameters are in marked contrast to those of the GlcNAc-P transferase that synthesizes GlcNAc-P-P-Und. When a similar kinetic analysis was performed in 5005Δ*gacI*, so that only GlcNAc-P-P-Und synthesis could be scored, an apparent Km for Und-P of 19.3 μM was obtained and the V_{max} was only 2.75 pmol/min/mg (Table 3, Fig. S6D). Clearly, GacI displays a much greater catalytic efficiency for the synthesis of GlcNAc-P-Und than the enzyme synthesizing GlcNAc-P-P-Und, which is reflected in the dramatic difference in the rates of synthesis of the two glycolipids. In summary, our data indicate that GacJ forms a stable association with GacI and stimulates the catalytic activity of GacI.

GlcNAc-P-P-Und is required for initiation of polyrhamnose polysaccharide biosynthesis. The GAS genome contains a close relative, *gacO*, of *E. coli wecA*, the tumicamycin-sensitive GlcNAc-phosphate transferase responsible for the synthesis of GlcNAc-P-P-Und in many bacteria (11, 12). Significantly, rhamnopolysaccharide synthesis in GAS is inhibited by tunicamycin (5) and a close homolog of GacO, *S. mutans* RpgP, has been shown to be required for rhamnopolysaccharide synthesis and can complement WecA deficiency in *E. coli* (13). Our analysis of UDP-[^{3}H]GlcNAc incorporation in GAS and GBS membrane lipids, revealing the presence of two GlcNAc-lipids, GlcNAc-P-Und and GlcNAc-P-P-Und, prompted us to investigate the role of GlcNAc-P-P-Und in the initiation of GAC biosynthesis. First, we investigated the role of GacO in the synthesis of GlcNAc-P-P-Und. When membrane fractions from the WecA deficient *E. coli* strain (CLM37) expressing GacO, CLM37:GacO, were incubated with UDP-[^{3}H]GlcNAc and Und-P (added as a dispersion in 1 % CHAPS), [^{3}H]GlcNAc-P-P-Und was formed at an enzymatic rate that is similar to that found in the WecA over-expressor strain, PR4019 (Fig. 7A). CLM37 carrying an empty vector synthesizes no detectable GlcNAc-P-P-Und under these conditions (Fig. 7A).

To test whether GlcNAc-P-P-Und might function as a membrane anchor for the synthesis of the polyrhamnose chain of the rhamnopolysaccharide, membrane fractions from MGAS5005 or the 5005Δ*gacI* mutant were pre-incubated with UDP-[^{3}H]GlcNAc and ATP, to form [^{3}H]GlcNAc-P-P-Und *in situ*, and chased with non-radioactive dTDP-rhamnose. The formation of the resultant [^{3}H]GlcNAc-lipids in 5005Δ*gacI* was analyzed by TLC on silica gel G after 30 min incubation as shown in (Fig. 7B and C). In the absence of TDP-rhamnose, 5005Δ*gacI* membranes produced only [^{3}H]GlcNAc-P-P-Und (Fig. 7B), whereas in the presence of TDP-rhamnose two radioactive products, [^{3}H]GlcNAc-P-P-Und and an additional product with slower mobility on TLC was observed (Fig 7C). Figure S7 shows the time course of accumulation of the new GlcNAc-lipid product. Significantly, parallel incubations with MGAS5005 showed that GlcNAc-P-Und was not glycosylated further. These results strongly support the possibility that GlcNAc-P-P-Und is an acceptor for rhamnosyl units in GAS and may function as the lipid anchor for polyrhamnosyl polysaccharide synthesis.

The GacL mutant displays increased sensitivity to peptidoglycan amidases. The Gram-positive cell wall protects interior structures, plasma membrane and peptidoglycan, from host defense peptides and hydrolytic and antimicrobial enzymes. It has been shown that the GacI mutant which is GlcNAc-deficient is more sensitive to LL-37-induced killing (5). To test the hypothesis that the loss of GlcNAc decorations in GAC alters cell wall permeability, we investigated the sensitivity of 5005Δ*gacL* to peptidoglycan amidases: PlyC (27), PlyPy (28) and CbpD (29). When 5005Δ*gacL* cells were grown in increasing concentrations of CbpD (Fig. 8A), PlyPy (Fig. 8B) or PlyC (Fig. 8C) cellular growth was dramatically inhibited compared to MGAS5005, indicating increased sensitivity to the presence of amidases.

DISCUSSION

In almost all Gram-positive bacteria, cell wall-attached glycopolymers are critical for cell envelope integrity and their depletion is lethal (30). Most streptococcal species including two human pathogens GAS and GBS do not synthesize WTA and instead produce Rha-containing glycopolymers as functional homologs of WTA (1). GAC is the major cell wall component of GAS and plays important roles in bacterial physiology and pathogenesis.

The polyrhamnose core of GAC is modified with GlcNAc in an approximately 2:1 ratio (3, 7) of rhamnose to GlcNAc. Collectively, the results of our study suggest a molecular mechanism of GAC biosynthesis in which rhamnan polymer is assembled at the cytoplasmic face of the plasma membrane, translocated to the cell surface and modified by GlcNAc on the outer side of the membrane as illustrated in Figure 9. We report that a lipid carrier, GlcNAc-P-P-Und, synthesized by GAS GacO, is a potential acceptor for initiation of rhamnan backbone biosynthesis. We speculate that the next step of rhamnan biosynthesis involves the action of the GacB, GacC, GacG and GacF glycosyltransferases. The lipid-anchored polyrhamnose is then translocated across the membrane by the ABC transporter encoded by *gacD* and *gacE*. This hypothesis is supported by studies of rhamnan biosynthesis in *S. mutans* (13, 31, 32).

Biosynthesis of the rhamnan-backbone of GAC is likely essential for GAS viability because the GacA enzyme involved in dTDP-rhamnose biosynthesis (14), and GacB and GacC glycosyltranferases are indispensable in GAS (5). In contrast, deletion of genes required for GlcNAc attachment to the rhamnan backbone does not affect GAS viability (5). This observation supports our hypothesis that biosynthesis of polyrhamnose and its translocation to the cell surface occur separately from the pathway involved in polyrhamnose modification with GlcNAc. In this study, we found that GlcNAc modification of rhamnan requires GlcNAc-P-Und synthesis. Previously, GlcNAc-P-Und was isolated from various *Bacilli* membranes (12, 23). However the enzyme required for GlcNAc-P-Und biosynthesis was not identified and the biological function of this lipid remained unknown. Our *in vitro* analysis of UDP-GlcNAc incorporation into GlcNAc-lipids by GAS membranes showed that GacI was required for the biosynthesis of GlcNAc-P-Und. Moreover CHAPS-soluble, affinity-purified GacI protein catalyzed the transfer of GlcNAc from UDP-GlcNAc to Und-P yielding GlcNac-P-Und in an *in vitro* reaction system.

Interestingly, we found that, in contrast to *B. cereus*, GAS and GBS membranes incubated with UDP-[^{3}H]GlcNAc synthesized primarily [^{3}H]GlcNAc-P-Und. Since GacI and GacO utilize a common pool of Und-P, this phenomenon is presumably due to a much higher apparent affinity of GacI for Und-P compared to GacO. Significantly, we observed a measurable increase in [^{3}H]GlcNAc incorporation into [^{3}H]GlcNAc-P-P-Und

in the GacI deletion mutant. This observation further confirmed that the relative synthetic rates of GlcNAc-P-Und and GlcNAc-P-P-Und biosynthesis are determined largely by Und-P availability. Furthermore, we found that the formation of GlcNAc-P-Und and GlcNAc-P-P-Und in GAS membranes was significantly stimulated by ATP. Since GacI activity is not stimulated directly by ATP, it is likely that the effect of ATP is due to increased formation of Und-P, *in situ*, via undecaprenol kinase activity. In *S. mutans,* membrane-associated undecaprenol kinase catalyzes the ATP-dependent phosphorylation of undecaprenol to Und-P (26). It is likely that the homolog of this enzyme encoded by M5005_Spy_0389 is responsible for Und-P biosynthesis in GAS.

Wzx flippases translocate Und-P-P-linked oligosaccharides from the cytoplasmic side to the periplasmic side of membranes in bacteria (33). We suggest that GacK may function to transport GlcNAc-P-Und to the extracellular space for utilization as GlcNAc donor in the GlcNAc modification of polyrhamnose. This hypothesis is based on the GlcNAc-deficient phenotype of the GacK mutant (5) and the high domain homology between GacK and proteins of the Wzx family of flippases. However, further research will be required to confirm this hypothesis. The phenotypes of the GacL mutant: absence of the GlcNAc side-chains in GAC and accumulation of GlcNAc-P-Und in the membrane is consistent with GacL functioning in the transfer of GlcNAc from GlcNAc-P-Und to polyrhamnose yielding GlcNAc-modified polysaccharide. Bioinformatics searches with the HHpred program against available protein structures identified the GT-C superfamily of glycosyltrasferases as the closest structural homologs of GacL. The members of the GT-C superfamily consist of proteins characterized with multiple transmembrane domains and a large periplasmic loop close to the N-terminus that contains a metal-binding motif (DxD, ExD, DDx or DEx) (34, 35). The GT-C enzymes act on the outside of the plasma membrane by transferring sugars from the lipid carrier, Und-P, onto either glycolipid or protein. GacL contains 12 transmembrane segments and the highly conserved DEX motif in the first extracellular loop. Thus, the predicted structural homology of GacL is in agreement with the proposed GacL function in GlcNAc transfer from GlcNAc-P-Und to polyrhamnose.

Thus, we suggest that the proposed mechanism of GlcNAc attachment to polyrhamnose in GAS is similar to the mechanism of O-antigen modification in *Shigella*

flexneri in which the GtrB glycosyltransferase, together with GtrA flippase and GtrV membrane protein, are required for addition of a glycosyl group to rhamnose of the O-antigen backbone polysaccharide, which is a part of lipopolysaccharide (36). The modification with glycosyl residues takes place after O-unit assembly and before transfer of the mature O-polysaccharide to the lipid A-core region of the lipopolysaccharide. Moreover, in *Mycobacterium tuberculosis* the modification of cell wall arabinogalactan with galactosamine occurs through similar mechanism involving N-acetyl galactosaminyl-phosphate-undecaprenol synthase, PpgS, and a membrane associated GT-C enzyme, Rv3779, that transfers a galactosamine residue from galactosaminyl-phosphate-undecaprenol to arabinogalactan (37).

The last step of GAC biosynthesis is probably similar to the last step of WTA biosynthesis: it involves attachment of GAC to certain MurNAc residues in peptidoglycan via a phosphate ester linkage (38). It is likely catalyzed by members of LytR-CpsA-Psr (LCP) phosphotransferase family encoded by M5005_Spy_1099 and M5005_Spy_1474.

GacI is predicted by the HHpred server to have structural homology to polyisoprenyl-glycosyltransferase GtrB from *Synechocystis* which is a homolog of *S. flexneri* GtrB (39). GtrB homologs in bacteria and eukaryotes belong to GT-A superfamily of glycosyltransferases and they are responsible for the addition of UDP-glucose to the bactoprenol carrier in the cytoplasm yielding the Und-P-glucose precursor in bacteria and dolichol phosphate-glucose precursor in eukaryotes (34, 35). Our BLAST search using the GacI sequence as query found homologs of this gene in many bacterial species, with the broadest diversity observed in the phylum Firmicutes (Supplemental File). To determine evolutionary relationships between the GacI homologs, these sequences were analyzed using CLANS (CLuster ANalysis of Sequences) (40) (Fig. 10). It is noteworthy that the *B. cereus* GacI has two homologs clustered in two different groups. Our bioinformatics analysis shows that these *gacI* homologs are located in different gene clusters encoding proteins involved in polysaccharide biosynthesis and transport. Moreover, the GAS GacI subgrouped with one *B. cereus* GacI. The GBS GacI and *E. faecalis* GacI homologs are located in the cluster well-separated from the GAS GacI subgroup. Thus, this observation points to

418 possible horizontal gene transfer between GAS and *Bacillus*. Analysis of UDP-GlcNAc
419 incorporation into GlcNAc-lipids in a GBS *gacI* knock-out strain confirmed the function of
420 this gene in the biosynthesis of GlcNAc-P-Und in GBS. Additionally, we demonstrated
421 that the *E. faecalis* GacI homolog (EpaI) functions in the biosynthesis of GlcNAc-P-Und.
422 Significantly, the presence of GacI homologs in GBS and *E. faecalis* matches the
423 reported occurrence of GlcNAc residues in the cell wall polysaccharides of these
424 bacteria (41-44).

425 GacJ, a small membrane protein with three transmembrane α-helices, is required
426 for GlcNAc side-chain attachment to polyrhamnose (5). GacJ belongs to DUF2304
427 family of proteins according to the Pfam protein family database (45). In *Geobacter* sp.,
428 *Desulfuromonas* sp., *Desulfurivibrio alkaliphilus, Desulfuromonas soudanensis* and
429 *Desulfurispirillum indicum*, DUF2304 homologous domains are fused with *gacI*
430 homologs (Fig. 10 and supplemental file). In *M. tuberculosis* a GacJ homolog, Rv3632,
431 is co-transcribed with the gene encoding N-acetyl galactosaminyl-phosphate-
432 undecaprenol synthase, PpgS, and is found to stimulate PpgS activity (37). Consistent
433 with this finding we showed that GacI and GacJ form a complex and co-expression of
434 GacI with GacJ significantly enhanced GacI catalytic activity. Further work is under way
435 to characterize the mechanism of GacJ action on GacI activity.

436 The identification of the mechanism of GlcNAc attachment to polyrhamnose raises
437 the question of how this modification functions biologically in bacteria. It has been
438 previously found that the Δ*gacI* mutant is hypersusceptible to human antimicrobial
439 peptide LL-37 and the antimicrobial action of factors released by thrombin-activated
440 platelets, suggesting a role of GlcNAc modification in protecting the plasma membrane
441 from antimicrobial agents (5). Our study identified the importance of GlcNAc for
442 protection of GAS peptidoglycan from amidase-induced lysis. In *E. faecalis* the GacI
443 homolog, EpaI, is involved in biosynthesis of a cell wall-attached polysaccharide (24),
444 however the polysaccharide structure of the mutant has not been elucidated. The Δ*epaI*
445 mutant was defective in conjugative transfer of a plasmid and resistance of bacteria to
446 detergent and bile salts (24).

447 In conclusion, our study provides a platform for elucidation of novel pathways of
448 glycopolymer biosynthesis in other bacterial pathogens including important drug-

resistant bacteria such as *E. faecalis* and *Clostridium sordellii* that possess GacI and GacJ homologs. Since the enzymes involved in the biosynthesis of cell wall attached glycopolymers are promising targets for novel antimicrobials and the glycopolymers represent important features for diagnostics and vaccine targets, our data may provide opportunities for developing novel therapeutics against antibiotic-resistant bacterial pathogens.

MATERIALS AND METHODS

Bacterial strains and growth conditions. All plasmids, strains and primers used in this study are listed in Tables S1 and S2 in the supplemental material. The strains used in this study were GAS M1-serotype strain MGAS5005 (16), *Bacillus cereus* (ATCC 14579), *Streptococcus agalactiae* COH1 (Group B *Streptococcus* or GBS), *E. coli* CLM37 (46), *E. coli* PR4019 (12), *E. coli* DH5α and *E. coli* Rosetta (DE3). GAS and GBS cultures were grown in Todd-Hewitt broth (BD) supplemented with 0.2% yeast extract (THY), or on THY agar plates at 37 °C. *E. coli* and *B. cereus* strains were grown in Luria-Bertani (LB) medium or on LB agar plates at 37 °C. When required, antibiotics were included at the following concentrations: ampicillin at 100 μg ml^{-1} for *E. coli*; streptomycin at 100 μg ml^{-1} for *E. coli*; erythromycin at 500 μg ml^{-1} for *E. coli* and 1 μg ml^{-1} for GAS and GBS; chloramphenicol at 10 μg ml^{-1} for *E. coli* and 5 μg ml^{-1} for GAS; spectinomycin at 200 μg ml^{-1} for *E. coli* and 100 μg ml^{-1} for GAS and GBS.

DNA techniques. Plasmid DNA was isolated from *E. coli* by commercial kits (Qiagen) according to the manufacturer's instructions and used to transform *E. coli*, GAS and GBS strains. Plasmids were transformed into GAS and GBS by electroporation as described previously (47). Chromosomal DNA was purified from GAS and GBS as described in (48). To construct single-base substitutions or deletion mutations, we used the QuikChange® II XL Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's protocol. Constructs containing mutations were identified by sequence analysis. Primers for site-directed mutagenesis are listed in Table S3. All constructs were confirmed by sequencing analysis (Eurofins MWG Operon).

Construction of the *gacI* deletion mutant in GAS. For construction of strain 5005Δ*gacI*, MGAS5005 chromosomal DNA was used as a template for amplification of

two DNA fragments using two primers pairs: GacIm-BamHI-f/GacIdel-r and GacIdel-f/GacIm-XhoI-r (Table S2). Primer GacIdel-f is complementary to primer GacIdel-r. The two gel-purified PCR products containing complementary ends were mixed and amplified using a PCR overlap method (49) with primer pair GacIm-BamHI-f/GacIm-XhoI-r to create the deletion of *gacI*. The PCR product was digested with BamHI and XhoI and ligated into BamHI/SalI-digested temperature-sensitive shuttle vector pJRS233 (50). The plasmid was designated pJRS233Δ*gacI* (Table S1). The resulting plasmid was transformed into MGAS5005, and erythromycin resistant colonies were selected on THY agar plates at 30 °C. Integration was performed by growth of transformants at 37 °C with erythromycin selection. Excision of the integrated plasmid was performed by serial passages in THY media at 30 °C and parallel screening for erythromycin-sensitive colonies. Mutants were verified by PCR sequencing of the loci.

Construction of the *gacL* deletion mutant in GAS. To create BglII and XhoI cloning sites in the polylinker region of pUC19, site-directed mutagenesis of the plasmid, using the primer pair listed in Table S3 was carried out. The plasmid was designated pUC19BX. The nonpolar *aadA* spectinomycin resistance cassette was amplified from pLR16T (Table S2) using primers Spec-SalI-f/Spec-BamH-r (Table S2), digested with SalI/BamHI, and ligated into SalI/BamHI-digested pUC19BX to yield pUC19BXspec. MGAS5005 chromosomal DNA was used as a template for amplification of two DNA fragments using two primers pairs: gacLup-BglII-f/gacLup-SalI-r and gacLdown-BamHI-f/gacLdown-XhoI-r (Table S2). The first PCR product was digested with BglII/SalI and ligated into BglII/SalI-digested pUC19BXspec. The resultant plasmid, pUC19BXspecL1, was digested with BamHI/XhoI and used for ligation with the second PCR product that was digested with BamHI/XhoI. The resultant plasmid, pUC19BXspecL2, was digested with BglII and XhoI to obtain a DNA fragment containing *aadA* flanked with the *gacL* upstream and downstream regions. The DNA fragment was ligated into pHY304 vector (51) and digested with BamHI/XhoI to yield pHY304Δ*gacL*. The resulting plasmid was transformed into MGAS5005, and erythromycin resistant colonies were selected on THY agar plates at 30 °C. The mutants were isolated as describe above. 5005Δ*gacL* mutants were screened for sensitivity to spectinomycin and verified by PCR sequencing of the loci.

Complementation of the 5005Δ*gacL* mutant with *gacL* (5005Δ*gacL gacL*$^{+}$). To construct the plasmid for complementation of the 5005Δ*gacL* mutant, MGAS5005 chromosomal DNA was used as a template for amplification of a wild-type copy of *gacL* using the primer pair GacL-XhoI-f and GacL-BglII-r (Table S2). The PCR products were digested with XhoI and BglII and cloned in pBBL740 (Table S1) previously digested with the respective enzymes. The integrational plasmid pBBL740 does not have a replication origin that is functional in GAS, so the plasmid can be maintained only by integrating into the GAS chromosome through homologous recombination. The resultant plasmid pGacL was transformed into 5005Δ*gacL* by electroporation and transformants were selected on agar plates containing chloramphenicol. Several chloramphenicol resistant colonies were selected and *gacL* integration into the chromosome was confirmed by sequencing a PCR fragment.

Construction of the *gacI* (SAN_1536) **deletion mutant in GBS.** GBS COH1 chromosomal DNA was used as a template for amplification of two DNA fragments using two primers pairs: gbsIup-BglII-f/gbsIup-SalI-r and gbsId-BamHI-f/gbsId-XhoI-r (Table S2). The plasmid for *gacI* (SAN_1536) knock-out was constructed using the same strategy described for *gacL* deletion (see above). The resultant plasmid, pHY304GBSΔI, was transformed into GBS COH1, and erythromycin resistant colonies were selected on THY agar plates at 30 °C. The mutants were isolated as described above. GBS COH1 Δ*gacI* mutants were screened for sensitivity to spectinomycin and verified by PCR sequencing of the loci.

Construction of the plasmids for *E. coli* expression of GacI, GacJ and GacO. To create a vector for expression of GacI from GAS, the gene was amplified from MGAS5005 chromosomal DNA using the primer pair GacI-NcoI-f and GacI-XhoI-r (Table S2). The PCR product was digested with NcoI and XhoI, and ligated into NcoII/XhoI-digested pRSF-NT vector (Table S1). The resultant plasmid, pGacI, contained *gacI* fused at the N-terminus with a His-tag followed by a TEV protease recognition site. To create a vector for expression of GacI and GacJ, the bicistronic DNA fragment was amplified from MGAS5005 chromosomal DNA using the primer pair GacI-NcoI-f and GacJ-XhoI-r (Table S2). The vector was constructed as described above. The plasmid was designated pGacIJ. pET21_NESG plasmid for expression of

the GacI homolog, EpaI, from *Enterococcus faecalis* was obtained from the DNASU repository (52). The construct was confirmed by sequencing analysis. The plasmids were transferred into competent *E. coli* Rosetta (DE3) (Novagen) using the manufacturer's protocol.

To create a vector for expression of GacO, the gene was amplified from MGAS5005 chromosomal DNA using the primer pair GacO-Xba-f and GacO-HindIII-r (Table S2). The PCR product was digested with XbaI and HindIII, and ligated into XbaI/HindIII-digested pBAD33 vector. The resultant plasmid, pGacO, was transferred into competent *E. coli* CLM37 strain that has a deletion of the *wecA* gene.

Construction of the plasmids for *E. coli* expression of PlyPy, CbpD and PlyC. To create a vector for expression of CbpD amidase (29), the gene was amplified from MGAS5005 chromosomal DNA using the primer pair 28-NcoI-f and 28-stop-r (Table S2). The PCR product was digested with NcoI and XhoI, and ligated into NcoI/XhoI-digested pRSF-NT vector. The resultant plasmid, pCbpD, contained *cbpD* fused at the N-terminus with a His-tag followed by a TEV protease recognition site. To create a vector for expression of PlyPy amidase (28), the gene was amplified from MGAS5005 chromosomal DNA using the primer pair PlyPy-NcoI-f and PlyPy-XhoI-r. The PCR product was digested with NcoI and XhoI, and ligated into NcoI/XhoI-digested pET-21d vector. The resultant plasmid, pPlyPy contained the gene fused at the C terminus with a His-tag.

To create a vector for expression of PlyC amidase (53), a DNA fragment spanning bicistronic operon which encodes *plyCA* and *plyCB* (27) was synthesized by ThermoFisher Scientific. The plasmid was digested with NcoI and XhoI, and ligated into NcoI/XhoI-digested pET21d vector. The resultant plasmid, pPlyC contained *plyCA* followed by *plyCB* fused at the C terminus with a His-tag.

Expression and purification of GacI, EpaI and GacI/GacJ complex. For expression of GacI, EpaI and GacI/GacJ complex, *E. coli* Rosetta (DE3) cells carrying the respective plasmid were grown to an OD_{600} of 0.4-0.6 and induced with 1 mM IPTG at 18 °C for approximately 16 hours. The cells were lysed in 20 mM Tris-HCl pH 7.5, 300 mM NaCl with two passes through a EmulsiFlex-C5 microfluidizer cell disrupter (Avestin, Inc., Ottawa Ontario,CA). The lysate was centrifuged at 7000 x g for 30

minutes, 4°C. The supernatant was centrifuged at 30,000 x g for 60 minutes, 4°C to isolate the membrane fraction.

To isolate GacI/GacJ complexes the membrane proteins were solubilized in 2.5% CHAPS, 20 mM Tris-HCl pH 7.5, 300 mM NaCl for 60 minutes, rotating at room temperature. Insoluble material was removed by centrifugation at 30,000 x g for 60 minutes, 4°C. Solubilized GacI/GacJ were purified by Ni-NTA chromatography with washes of 2.5% CHAPS, 20 mM Tris-HCl pH 7.5, 300 mM NaCl and 2.5% CHAPS, 20 mM Tris-HCl pH7.5, 300 mM NaCl, 10 mM imidazole, and elution with 2.5% CHAPS, 20 mM Tris-HCl pH 7.5, 300 mM NaCl, 250 mM imidazole. The GacI/GacJ complex was further purified by size exclusion chromatography on a Superdex 200 (GE Biosciences) column in 2.5% CHAPS, 20 mM HEPES pH 7.5, 100 mM NaCl, with monitoring for protein elution at 280 nm.

Expression of GacO. For expression of GacO, *E. coli* CLM37 cells carrying the pertinent plasmid were grown to an OD_{600} of 0.8 and induced with 13 mM L-arabinose at 25 °C for approximately 3 hours. The cells were lysed in 20 mM Tris-HCl pH 7.5, 300 mM NaCl with two passes through a microfluidizer cell disrupter. The lysate was centrifuged at 1000 x g for 15 minutes, 4 °C. The supernatant was centrifuged at 40,000 x g for 60 minutes, 4 °C to isolate the membrane fraction.

Expression and purification of PlyPy, CbpD and PlyC. For expression and purification of PlyPy, CbpD and PlyC, *E. coli* Rosetta (DE3) cells carrying the respective plasmids were grown to an OD_{600} of 0.4-0.6 and induced at 18 °C with 1 mM IPTG for approximately 16 hours. The cells were lysed in 20 mM Tris-HCl pH 7.5, 300 mM NaCl with two passes through a microfluidizer cell disrupter. The soluble fraction was purified by Ni-NTA chromatography with washes of 20 mM Tris-HCl pH 7.5, 300 mM NaCl and 20 mM Tris-HCl pH7.5, 300 mM NaCl, 10mM imidazole, and elution with 20 mM Tris-HCl pH 7.5, 300 mM NaCl, 250 mM imidazole. The PlyPy and CbpD eluate was further purified by size exclusion chromatography on a Superdex 200 column in 20 mM HEPES pH 7.5, 100 mM NaCl for PlyPy, or 20 mM MOPS pH 6.5, 100 mM NaCl for CbpD. PlyC was further purified by anion exchange chromatography on a MonoQ 5/50 GL (GE Biosciences) column in 10 mM sodium phosphate pH 6.0, and a 20 column volume elution gradient of 0-500 mM NaCl.

Isolation of GAS and GBS membranes. Bacteria were grown at 37 °C to an OD_{600} of 0.8. To obtain GAS cell membranes, cell pellet was re-suspended in phosphate-buffered saline (PBS) and incubated 1 h with PlyC lysin as described in (54). To obtain GBS membranes, the cell pellet was re-suspended in acetate buffer (50 mM CH_3COONa, 10 mM CaCl, 50 mM NaCl, pH 5) and incubated with mutanolysin (200 U/ml, Sigma–Aldrich) for 2 h at 37 °C. After hydrolytic enzyme treatment the bacterial suspension was sonicated using a Fisher Scientific™ Model 505 Sonic Dismembrator, 15 times ×15 s. After centrifugation at 8,000 *g* for 10 min at 4 °C, the supernatant was collected and centrifuged at 40,000 *g* for 60 min. The pellet was collected as the membrane fraction.

Mass-spectrometry analysis of GacI and GacJ. LC-MS/MS for proteomic analysis was performed using an LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific) coupled with an Eksigent Nanoflex cHiPLC system (Eksigent) through a nanoelectrospray ionization source. The LC-MS/MS data were subjected to database searches for protein identification using Proteome Discoverer software V. 1.3 (Thermo Fisher Scientific) with a local MASCOT search engine.

Dot-blot analysis. Bacterial cells (1 ml) from exponential phase cultures (OD_{600}=0.8) was centrifuged, washed with PBS, resuspended in 100 µl of PBS and incubated with 1 µl PlyC (1.5 mg/ml) for 1 h at 37°C. After centrifugation at 16,000 *g* for 2 min, 5 µl of the supernatant was spotted on a nitrocellulose membrane. The membrane was blocked 1 h with 7% skim milk in PBS with 0.1% Tween 20, and incubated overnight with an anti-GAC antibody diluted 1:5,000 (ab9191; Abcam). Bound antibodies were detected with a peroxidase-conjugated goat anti-rabbit immunoglobulin G antibody, and the Amersham ECL (enhanced chemiluminescence) Western blotting system.

Binding of succinylated wheat germ agglutinin (sWGA) to bacteria. MGAS5005 and 5005Δ*gacL* cells were collected during mid-exponential phase (OD_{600}=0.6), washed three times with BSA-saline solution (0.5% BSA, 0.15 M NaCl), resuspended in BSA-saline solution and incubated for 30 minute at 37 °C with mixing. GlcNAc specific fluorescein-sWGA was added to final concentrations of 0, 12.5, 25, 50, 100 µg/ml. After 1 h of incubation at 37 °C with mixing, the cells were centrifuged, washed twice, and

resuspended in BSA-saline. Bound fluorescein-sWGA was quantified in a fluorimeter SpectraMax M5 (Molecular Devices) using an excitation of 544 nm and emission of 590 nm.

Isolation of cell wall from GAS. MGAS5005 and 5005Δ*gacL* cell wall was prepared from exponential phase cultures (OD_{600}=0.8) by the SDS-boiling procedure as described for *Streptococcus pneumoniae* (55). Purified cell wall samples were lyophilized and used for composition analysis.

Carbohydrate composition analysis. Carbohydrate composition analysis was performed at the Complex Carbohydrate Research Center (Athens, GA) by combined gas chromatography/mass spectrometry (GC/MS) of the per-O-trimethylsilyl (TMS) derivatives of the monosaccharide methyl glycosides produced from the sample by acidic methanolysis as described previously by (56).

Assay of lytic activity of PlyPy, CbpD and PlyC. Aliquots of MGAS5005 and 5005Δ*gacL* frozen stocks [prepared as described in (57)] were inoculated into THY medium 1:20. To a 96 well plate, 200 µL aliquots of culture were dispensed in duplicate, per condition. The plates were grown at 37 °C without aeration for 2 hours. An assay plate was set up containing PlyPy (8 µg mL^{-1}, 16 µg mL^{-1}, 32 µg mL^{-1}), CbpD (0.31 µg mL^{-1}, 0.63 µg mL^{-1}, 1.26 µg mL^{-1}) or PlyC (1.6 ng mL^{-1}, 3.1 ng mL^{-1}, 6.2 ng mL^{-1}) in a volume of 100 µL. To the assay plate, 100 µL of culture was added to each well giving final concentrations of: PlyPy (4 µg mL^{-1}, 8 µg mL^{-1}, 16 µg mL^{-1}), CbpD (0.16 µg mL^{-1}, 0.31 µg mL^{-1}, 0.63 µg mL^{-1}), or PlyC (0.8 ng mL^{-1}, 1.6 ng mL^{-1}, 3.1 ng mL^{-1}). After mixing, the absorbance at 600 nm was measured for each well (t_0) and the plate was incubated at 37 °C for 1 h. The absorbance at 600 nm was then measured again (t_1). The percentage difference (t_1-t_0) in growth relative to 0 mM of enzyme/antimicrobial, for each concentration was calculated (i.e. 0 mM PlyPy = 100%).

Extraction and characterization of bacterial glycolipids. Two liters of bacterial cells from exponential phase cultures (OD_{600}=0.8) were recovered by sedimentation at 10,000 x g for 30 min, and washed with ice-cold PBS 2x. Cells were resuspended in 50 ml PBS, sensitized by incubation with 10 ng mL^{-1} PlyC for 1 h at 37 °C and stirred vigorously with two volumes of CH_3OH and one volume of $CHCl_3$ for 30 min at room temperature. The mixture was divided into 5 mL aliquots in 12 mL glass centrifuge

tubes. Insoluble material was removed by centrifugation at 200 *g* and the organic extract was transferred to a separatory funnel. The insoluble residue was further extracted with 1 mL of $CHCl_3/CH_3OH$ (2:1) per tube, two times, and the extracts were combined with the previous organic phase. The organic extract was supplemented with $CHCl_3$ and 0.9% NaCl to give a final composition of $CHCl_3/CH_3OH$/0.9% NaCl (3:2:1), mixed vigorously, and allowed to stand in the cold until phase separation was achieved. The organic phase was drained off into a second separatory funnel and the organic layer was washed with 1/3 volume of $CHCl_3/CH_3OH$/0.9% saline (3:48:47), two times. The aqueous layers were discarded. The organic extract was dried on a vacuum rotary evaporator, dissolved in a small volume of $CHCl_3/CH_3OH$ (2:1) and transferred to a 12.5 x 100 mm screw cap glass tube (with Teflon lined cap). The organic extract was dried under a stream of nitrogen gas and the glycerolipids were destroyed by deacylation in 0.1 M KOH in toluene/CH_3OH (1:3) at 0 °C, 60 min. Following deacylation, the reactions were neutralized with acetic acid, diluted with two volumes of $CHCl_3$, 1 volume of $CHCl_3/CH_3OH$ (2:1) and 1/5 volume of 0.9% NaCl/10 mM EDTA. The two-phase mixture was mixed vigorously and centrifuged to separate the phases. The organic phase was dried under nitrogen, spotted on a 20 x 20 cm sheet of Silica Gel G and developed in $CHCl_3/CH_3OH/H_2O/NH_4OH$ (65:25:4:1). Bacterial lipids were visualized by staining with iodine vapors and pertinent spots were scraped from the thin layer plate, eluted from the silica gel with $CHCl_3/CH_3OH$ (2:1) and reserved for further analysis.

Mass spectrometry analysis of a phospholipid isolated from 5005Δ*gacL*. A phospholipid accumulated by 5005Δ*gacL* was purified as described above using preparative TLC in silica gel. The compound was analyzed by LC-MS using a Q-exactive mass spectrometer and an Ultimate 3000 ultra high performance liquid chromatography system (Thermo Fisher Scientific, San Jose, CA) on a Kinetex C18 reversed-phase column (2.6 mm × 100 mm, 2.1 µm, Phenomenex, USA). Two solvents were used for gradient elution: (A) acetonitrile/water (2:3, v/v), (B) isopropanol/acetonitrile (9:1, v/v). Both A and B contained 10 mM ammonium formate and 0.1% formic acid. The column temperature was maintained at 40 °C, and the flow rate was set to 0.25 ml/min. Mass spectrometric detection was performed by electrospray ionization in negative ionization mode with source voltage maintained at

4.0 kV. The capillary temperature, sheath gas flow and auxiliary gas flow were set at 330 °C, 35 arb and 12 arb, respectively. Full-scan MS spectra (mass range m/z 400 to 1500) were acquired with resolution R = 70,000 and AGC target 5e5. MS/MS fragmentation was performed using high-energy C-trap dissociation with resolution R = 35,000 and AGC target 1e6. The normalized collision energy was set at 30.

Assay for incorporation of GlcNAc into polyisoprenol-linked glycolipid intermediates. Reaction mixtures for measuring the incorporation of [^{3}H]GlcNAc into lipids contained 50 mM Tris, pH 7.4, 0.25 M sucrose, 20 mM $MgCl_2$, 0.5 mM β-mercaptoethanol, 5 μM UDP-[6-^{3}H]GlcNAc [(100-2000 cpm/pmol) American Radiolabelled Chemicals] and bacterial membrane suspension (50-250 μg bacterial membrane protein) in a total volume of 10 to 100 μL. In some experiments, 1 mM ATP was included, as indicated. Following incubation at 30 °C, the enzymatic reactions were terminated by the addition of 40 volumes of $CHCl_3$/CH_3OH (2:1), thoroughly mixed and incubated for 5 min at room temperature. Insoluble material was sedimented at 200 *g* and the organic extract was transferred to a 12x100 mm glass tube. The residue was re-extracted with 1 ml $CHCl_3$/CH_3OH (2:1), two times, and the organic extracts were combined. The pooled organic extracts were freed of unincorporated radioactivity by sequential partitioning with 1/5 volume of 0.9% saline/10 mM EDTA and then with 1/3 volume of $CHCl_3$/CH_3OH/0.9% saline (3:48:47) 3 times, discarding the aqueous phase each time. The washed organic phases were dried under nitrogen and re-dissolved in $CHCl_3$/CH_3OH (2:1). A carefully measured aliquot was removed and analyzed for radioactivity by liquid scintillation spectrometry after drying. The remainder of the sample was analyzed by thin layer chromatography on Silica Gel G, developed in $CHCl_3$/CH_3OH/H_2O/NH_4OH (65:25:4:1), using a BioScan AR2000 radiochromatoscanner. Incorporation of [^{3}H]GlcNAc into individual [^{3}H]GlcNAc-lipids was calculated using the peak integration values obtained from the BioScan.

Degradation of GlcNAc-lipids by mild acid and mild alkaline treatment. Bacterial lipids were subjected to mild acid hydrolysis (50% isopropanol, 0.1 M HCl, 50 °C, 1 h) and mild alkaline methanolysis (0.1 M KOH in toluene/CH_3OH (1:3), 0 °C, 1 h). Following incubation, the reactions were neutralized with either 1 M Tris or concentrated acetic acid, diluted with $CHCl_3$, CH_3OH and 0.9% NaCl to give a final composition of

728 3:2:1, respectively and partitioned as described above. The organic phases were dried
729 under nitrogen and either quantified for radioactivity or analyzed by thin layer
730 chromatography and detected by iodine staining.

731 **Phenolysis of GlcNAc-P-Und.** Purified [^{3}H]GlcNAc-P-Und was dried under
732 nitrogen in a 12x100 mm conical screw-cap tube and heated to 68 °C in 0.2 ml 50%
733 aqueous phenol as described by Murazumi et al. (22). Following phenolysis, 0.1 ml
734 water was added and the samples were thoroughly mixed. The aqueous and phenolic
735 layers were separated, dried and analyzed for radioactivity by scintillation spectrometry.

736 **Analytical methods.** Membrane protein concentrations were determined using the
737 BCA protein assay kit (Pierce Chemical Co.) employing the method of Ruiz et al. (33).
738 Radioactivity was quantified by liquid scintillation spectrometry on a Packard Tri-Carb
739 Liquid Scintillation Spectrometer using Econosafe Complete Counting Cocktail
740 (Research Products International, Inc., Elk Grove IL).

741 **Phylogenetic analysis of GacI homologs.** Sequences of GacI homologs were
742 retrieved from the non-redundant database using BLAST (58). One thousand
743 sequences were downloaded and were manually curated to remove sequences with
744 shorter that 90% sequence length overlap. The redundancy was reduced using CD-HIT
745 with 0.98 cut off (59). Sequences were manually curated to correct misannotated
746 translation start codons. The pair-wise similarities were analyzed and visualized using
747 CLANS with an E-value cut off 1 e^{-80} (40, 60, 61).

748 **Bioinformatics tools.** The TOPCONS (http://topcons.net/) (62) web server was
749 employed to predict trans-membrane regions of GacI, GacJ and GacL. Homology
750 detection and structure prediction were performed by the HHpred server
751 (https://toolkit.tuebingen.mpg.de/#/tools/hhpred) (63).

752 **Statistical Analysis.** Unless otherwise indicated, statistical analysis was carried out
753 from at least three independent experiments. Quantitative data was analyzed using the
754 paired Student's t-test. A P value equal to or less that 0.05 is considered statistically
755 significant.

756

757 **Supplemental Material**

Additional Supplemental Material may be found in the online version of this article at the publisher's web-site.

Table S1

Table S2

Table S3

Table S4

Figure S1

Figure S2

Figure S3

Figure S4

Figure S5

Figure S6

Figure S7

Supplemental File 1

ACKNOWLEDGEMENTS

J.S.R., R.J.E., H.Z., A.J.M., N.M.vS., K.V.K. and N.K. designed the experiments. J.S.R., R.J.E. and N.K. performed biochemical experiments. P.D., J.C., H.Z. and A.J.M. performed MS analysis. N.K. constructed plasmids and isolated mutants. J.S.R., R.J.E., K.V.K. and N.K. wrote the manuscript.

Funding Information

This work was supported by NIH grants R21AI113253 from the National Institute of Allergy and Infectious Diseases (to N.K.) and R01GM102129 (to Dr. Charles J. Waechter and J.S.R), 1S10OD021753 to A.J.M. and by the Center of Biomedical Research Excellence (COBRE) Pilot Grant (to K.V.K, N.K. and J.S.R.) supported by NIH grant P30GM110787 from the National Institute of General Medical Sciences.

784 N.M.vS. is supported by VIDI grant 91713303 from the Netherlands Organization for
785 Scientific Research (NWO).

786 Carbohydrate composition analysis at the Complex Carbohydrate Research Center
787 was supported by the Chemical Sciences, Geosciences and Biosciences Division,
788 Office of Basic Energy Sciences, U.S. Department of Energy grant (DE-FG02-
789 93ER20097) to Parastoo Azadi.

790 The funders had no role in study design, data collection and interpretation, or the
791 decision to submit the work for publication

792

793

REFERENCES

1. **Mistou MY, Sutcliffe IC, van Sorge NM.** 2016. Bacterial glycobiology: rhamnose-containing cell wall polysaccharides in Gram-positive bacteria. FEMS Microbiol Rev **40:**464-479.

2. **Carapetis JR, Steer AC, Mulholland EK, Weber M.** 2005. The global burden of group A streptococcal diseases. Lancet Infect Dis **5:**685-694.

3. **McCarty M.** 1952. The lysis of group A hemolytic streptococci by extracellular enzymes of Streptomyces albus. II. Nature of the cellular substrate attacked by the lytic enzymes. J Exp Med **96:**569-580.

4. **Lancefield RC.** 1933. A Serological Differentiation of Human and Other Groups of Hemolytic Streptococci. J Exp Med **57:**571-595.

5. **van Sorge NM, Cole JN, Kuipers K, Henningham A, Aziz RK, Kasirer-Friede A, Lin L, Berends ET, Davies MR, Dougan G, Zhang F, Dahesh S, Shaw L, Gin J, Cunningham M, Merriman JA, Hutter J, Lepenies B, Rooijakkers SH, Malley R, Walker MJ, Shattil SJ, Schlievert PM, Choudhury B, Nizet V.** 2014. The classical lancefield antigen of group a Streptococcus is a virulence determinant with implications for vaccine design. Cell host & microbe **15:**729-740.

6. **Coligan JE, Kindt TJ, Krause RM.** 1978. Structure of the streptococcal groups A, A-variant and C carbohydrates. Immunochemistry **15:**755-760.

7. **McCarty M.** 1956. Variation in the group-specific carbohydrate of group A streptococci. II. Studies on the chemical basis for serological specificity of the carbohydrates. J Exp Med **104:**629-643.

817 8. **Kabanova A, Margarit I, Berti F, Romano MR, Grandi G, Bensi G, Chiarot E,**
818 **Proietti D, Swennen E, Cappelletti E, Fontani P, Casini D, Adamo R, Pinto V,**
819 **Skibinski D, Capo S, Buffi G, Gallotta M, Christ WJ, Campbell AS, Pena J,**
820 **Seeberger PH, Rappuoli R, Costantino P.** 2010. Evaluation of a Group A
821 Streptococcus synthetic oligosaccharide as vaccine candidate. Vaccine **29:**104-
822 114.
823 9. **Sabharwal H, Michon F, Nelson D, Dong W, Fuchs K, Manjarrez RC, Sarkar**
824 **A, Uitz C, Viteri-Jackson A, Suarez RS, Blake M, Zabriskie JB.** 2006. Group A
825 streptococcus (GAS) carbohydrate as an immunogen for protection against GAS
826 infection. J Infect Dis **193:**129-135.
827 10. **Soldo B, Lazarevic V, Karamata D.** 2002. tagO is involved in the synthesis of all
828 anionic cell-wall polymers in Bacillus subtilis 168. Microbiology **148:**2079-2087.
829 11. **Meier-Dieter U, Barr K, Starman R, Hatch L, Rick PD.** 1992. Nucleotide
830 sequence of the Escherichia coli rfe gene involved in the synthesis of
831 enterobacterial common antigen. Molecular cloning of the rfe-rff gene cluster. J
832 Biol Chem **267:**746-753.
833 12. **Rush JS, Rick PD, Waechter CJ.** 1997. Polyisoprenyl phosphate specificity of
834 UDP-GlcNAc:undecaprenyl phosphate N-acetylglucosaminyl 1-P transferase
835 from E.coli. Glycobiology **7:**315-322.
836 13. **Yamashita Y, Shibata Y, Nakano Y, Tsuda H, Kido N, Ohta M, Koga T.** 1999.
837 A novel gene required for rhamnose-glucose polysaccharide synthesis in
838 Streptococcus mutans. J Bacteriol **181:**6556-6559.

14. **van der Beek SL, Le Breton Y, Ferenbach AT, Chapman RN, van Aalten DM, Navratilova I, Boons GJ, McIver KS, van Sorge NM, Dorfmueller HC.** 2015. GacA is essential for Group A Streptococcus and defines a new class of monomeric dTDP-4-dehydrorhamnose reductases (RmlD). Mol Microbiol **98:**946-962.

15. **Shibata Y, Yamashita Y, Ozaki K, Nakano Y, Koga T.** 2002. Expression and characterization of streptococcal rgp genes required for rhamnan synthesis in Escherichia coli. Infect Immun **70:**2891-2898.

16. **Sumby P, Porcella SF, Madrigal AG, Barbian KD, Virtaneva K, Ricklefs SM, Sturdevant DE, Graham MR, Vuopio-Varkila J, Hoe NP, Musser JM.** 2005. Evolutionary origin and emergence of a highly successful clone of serotype M1 group a Streptococcus involved multiple horizontal gene transfer events. J Infect Dis **192:**771-782.

17. **Bevenue A, Williams KT.** 1951. Further evidence indicating the specificity of the orcinol spray reagent for ketoheptoses on paper chromatograms. Arch Biochem Biophys **34:**225-227.

18. **Dittmer JC, Lester RL.** 1964. A Simple, Specific Spray for the Detection of Phospholipids on Thin-Layer Chromatograms. Journal of lipid research **5:**126-127.

19. **Danilov LL, Druzhinina TN, Kalinchuk NA, Maltsev SD, Shibaev VN.** 1989. Polyprenyl phosphates: synthesis and structure-activity relationship for a biosynthetic system of Salmonella anatum O-specific polysaccharide. Chemistry and physics of lipids **51:**191-203.

862 20. **Wolucka BA, de Hoffmann E, Rush JS, Waechter CJ.** 1996. Determination of
863 the anomeric configuration of glycosyl esters of nucleoside pyrophosphates and
864 polyisoprenyl phosphates by fast-atom bombardment tandem mass
865 spectrometry. Journal of the American Society for Mass Spectrometry **7:**541-549.
866 21. **Wolucka BA, Rush JS, Waechter CJ, Shibaev VN, de Hoffmann E.** 1998. An
867 electrospray-ionization tandem mass spectrometry method for determination of
868 the anomeric configuration of glycosyl 1-phosphate derivatives. Analytical
869 biochemistry **255:**244-251.
870 22. **Murazumi N, Yamamori S, Araki Y, Ito E.** 1979. Anomeric configuration of N-
871 acetylglucosaminyl phosphorylundecaprenols formed in Bacillus cereus
872 Membranes. J Biol Chem **254:**11791-11793.
873 23. **Yamamori S, Murazumi N, Araki Y, Ito E.** 1978. Formation and function of N-
874 acetyloglucosamine-linked phosphoryl- and pyrophosphorylundecaprenols in
875 membranes from Bacillus cereus. J Biol Chem **253:**6516-6522.
876 24. **Dale JL, Cagnazzo J, Phan CQ, Barnes AM, Dunny GM.** 2015. Multiple roles
877 for Enterococcus faecalis glycosyltransferases in biofilm-associated antibiotic
878 resistance, cell envelope integrity, and conjugative transfer. Antimicrob Agents
879 Chemother **59:**4094-4105.
880 25. **Reusch VM, Jr., Panos C.** 1977. Synthesis of "group polysaccharide" by
881 membranes from Streptococcus pyogenes and its stabilized L-form. J Bacteriol
882 **129:**1407-1414.

883 26. **Lis M, Kuramitsu HK.** 2003. The stress-responsive dgk gene from
884 Streptococcus mutans encodes a putative undecaprenol kinase activity. Infect
885 Immun **71:**1938-1943.
886 27. **Nelson D, Schuch R, Chahales P, Zhu S, Fischetti VA.** 2006. PlyC: a
887 multimeric bacteriophage lysin. Proc Natl Acad Sci U S A **103:**10765-10770.
888 28. **Lood R, Raz A, Molina H, Euler CW, Fischetti VA.** 2014. A highly active and
889 negatively charged Streptococcus pyogenes lysin with a rare D-alanyl-L-alanine
890 endopeptidase activity protects mice against streptococcal bacteremia.
891 Antimicrob Agents Chemother **58:**3073-3084.
892 29. **Berg KH, Ohnstad HS, Havarstein LS.** 2012. LytF, a novel competence-
893 regulated murein hydrolase in the genus Streptococcus. J Bacteriol **194:**627-635.
894 30. **Chu M, Mallozzi MJ, Roxas BP, Bertolo L, Monteiro MA, Agellon A,**
895 **Viswanathan VK, Vedantam G.** 2016. A Clostridium difficile Cell Wall
896 Glycopolymer Locus Influences Bacterial Shape, Polysaccharide Production and
897 Virulence. PLoS Pathog **12:**e1005946.
898 31. **Tsukioka Y, Yamashita Y, Nakano Y, Oho T, Koga T.** 1997. Identification of a
899 fourth gene involved in dTDP-rhamnose synthesis in Streptococcus mutans. J
900 Bacteriol **179:**4411-4414.
901 32. **Yamashita Y, Tsukioka Y, Tomihisa K, Nakano Y, Koga T.** 1998. Genes
902 involved in cell wall localization and side chain formation of rhamnose-glucose
903 polysaccharide in Streptococcus mutans. J Bacteriol **180:**5803-5807.
904 33. **Ruiz N.** 2015. Lipid Flippases for Bacterial Peptidoglycan Biosynthesis. Lipid
905 insights **8:**21-31.

34. **Gomez-Casati DF, Martin M, Busi MV.** 2013. Polysaccharide-synthesizing glycosyltransferases and carbohydrate binding modules: the case of starch synthase III. Protein and peptide letters **20:**856-863.

35. **Lairson LL, Henrissat B, Davies GJ, Withers SG.** 2008. Glycosyltransferases: structures, functions, and mechanisms. Annu Rev Biochem **77:**521-555.

36. **Korres H, Mavris M, Morona R, Manning PA, Verma NK.** 2005. Topological analysis of GtrA and GtrB proteins encoded by the serotype-converting cassette of Shigella flexneri. Biochem Biophys Res Commun **328:**1252-1260.

37. **Skovierova H, Larrouy-Maumus G, Pham H, Belanova M, Barilone N, Dasgupta A, Mikusova K, Gicquel B, Gilleron M, Brennan PJ, Puzo G, Nigou J, Jackson M.** 2010. Biosynthetic origin of the galactosamine substituent of Arabinogalactan in Mycobacterium tuberculosis. J Biol Chem **285:**41348-41355.

38. **Schaefer K, Matano LM, Qiao Y, Kahne D, Walker S.** 2017. In vitro reconstitution demonstrates the cell wall ligase activity of LCP proteins. Nature chemical biology **13:**396-401.

39. **Ardiccioni C, Clarke OB, Tomasek D, Issa HA, von Alpen DC, Pond HL, Banerjee S, Rajashankar KR, Liu Q, Guan Z, Li C, Kloss B, Bruni R, Kloppmann E, Rost B, Manzini MC, Shapiro L, Mancia F.** 2016. Structure of the polyisoprenyl-phosphate glycosyltransferase GtrB and insights into the mechanism of catalysis. Nat Commun **7:**10175.

40. **Frickey T, Lupas A.** 2004. CLANS: a Java application for visualizing protein families based on pairwise similarity. Bioinformatics **20:**3702-3704.

928 41. **Hancock LE, Gilmore MS.** 2002. The capsular polysaccharide of Enterococcus
929 faecalis and its relationship to other polysaccharides in the cell wall. Proc Natl
930 Acad Sci U S A **99:**1574-1579.
931 42. **Michon F, Chalifour R, Feldman R, Wessels M, Kasper DL, Gamian A,**
932 **Pozsgay V, Jennings HJ.** 1991. The alpha-L-(1----2)-trirhamnopyranoside
933 epitope on the group-specific polysaccharide of group B streptococci. Infect
934 Immun **59:**1690-1696.
935 43. **Pritchard DG, Gray BM, Dillon HC, Jr.** 1984. Characterization of the group-
936 specific polysaccharide of group B Streptococcus. Arch Biochem Biophys
937 **235:**385-392.
938 44. **Teng F, Singh KV, Bourgogne A, Zeng J, Murray BE.** 2009. Further
939 characterization of the epa gene cluster and Epa polysaccharides of
940 Enterococcus faecalis. Infect Immun **77:**3759-3767.
941 45. **Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC,**
942 **Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A.**
943 2016. The Pfam protein families database: towards a more sustainable future.
944 Nucleic Acids Res **44:**D279-285.
945 46. **Linton D, Dorrell N, Hitchen PG, Amber S, Karlyshev AV, Morris HR, Dell A,**
946 **Valvano MA, Aebi M, Wren BW.** 2005. Functional analysis of the
947 Campylobacter jejuni N-linked protein glycosylation pathway. Mol Microbiol
948 **55:**1695-1703.
949 47. **Hoff JS, DeWald M, Moseley SL, Collins CM, Voyich JM.** 2011. SpyA, a C3-
950 like ADP-ribosyltransferase, contributes to virulence in a mouse subcutaneous

951 model of Streptococcus pyogenes infection. Infection and immunity **79:**2404-
952 2411.
953 48. **Caparon MG, Scott JR.** 1991. Genetic manipulation of pathogenic streptococci.
954 Methods Enzymol **204:**556-586.
955 49. **Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR.** 1989. Engineering hybrid
956 genes without the use of restriction enzymes: gene splicing by overlap extension.
957 Gene **77:**61-68.
958 50. **Ruiz N, Wang B, Pentland A, Caparon M.** 1998. Streptolysin O and adherence
959 synergistically modulate proinflammatory responses of keratinocytes to group A
960 streptococci. Mol Microbiol **27:**337-346.
961 51. **Chaffin DO, Beres SB, Yim HH, Rubens CE.** 2000. The serotype of type Ia and
962 III group B streptococci is determined by the polymerase gene within the
963 polycistronic capsule operon. J Bacteriol **182:**4466-4477.
964 52. **Seiler CY, Park JG, Sharma A, Hunter P, Surapaneni P, Sedillo C, Field J,**
965 **Algar R, Price A, Steel J, Throop A, Fiacco M, LaBaer J.** 2014. DNASU
966 plasmid and PSI:Biology-Materials repositories: resources to accelerate
967 biological research. Nucleic Acids Res **42:**D1253-1260.
968 53. **Nelson D, Loomis L, Fischetti VA.** 2001. Prevention and elimination of upper
969 respiratory colonization of mice by group A streptococci by using a bacteriophage
970 lytic enzyme. Proc Natl Acad Sci U S A **98:**4107-4112.
971 54. **Raz A, Fischetti VA.** 2008. Sortase A localizes to distinct foci on the
972 Streptococcus pyogenes membrane. Proc Natl Acad Sci U S A **105:**18549-
973 18554.

974 55. **Bui NK, Eberhardt A, Vollmer D, Kern T, Bougault C, Tomasz A, Simorre JP,**
975 **Vollmer W.** 2012. Isolation and analysis of cell wall components from
976 Streptococcus pneumoniae. Analytical biochemistry **421:**657-666.
977 56. **Santander J, Martin T, Loh A, Pohlenz C, Gatlin DM, 3rd, Curtiss R, 3rd.**
978 2013. Mechanisms of intrinsic resistance to antimicrobial peptides of
979 Edwardsiella ictaluri and its influence on fish gut inflammation and virulence.
980 Microbiology **159:**1471-1486.
981 57. **Edgar RJ, Chen J, Kant S, Rechkina E, Rush JS, Forsberg LS, Jaehrig B,**
982 **Azadi P, Tchesnokova V, Sokurenko EV, Zhu H, Korotkov KV, Pancholi V,**
983 **Korotkova N.** 2016. SpyB, a Small Heme-Binding Protein, Affects the
984 Composition of the Cell Wall in Streptococcus pyogenes. Frontiers in cellular and
985 infection microbiology **6:**126.
986 58. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local
987 alignment search tool. J Mol Biol **215:**403-410.
988 59. **Huang Y, Niu B, Gao Y, Fu L, Li W.** 2010. CD-HIT Suite: a web server for
989 clustering and comparing biological sequences. Bioinformatics **26:**680-682.
990 60. **Alva V, Nam SZ, Soding J, Lupas AN.** 2016. The MPI bioinformatics Toolkit as
991 an integrative platform for advanced protein sequence and structure analysis.
992 Nucleic Acids Res **44:**W410-415.
993 61. **Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K,**
994 **Madden TL.** 2009. BLAST+: architecture and applications. BMC bioinformatics
995 **10:**421.

62. **Tsirigos KD, Peters C, Shu N, Kall L, Elofsson A.** 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. Nucleic Acids Res **43:**W401-407.

63. **Soding J, Biegert A, Lupas AN.** 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res **33:**W244-248.

1003 # Tables

1004 Table 1. The major fragments of the precursor ion at *m/z* 1048

m/z	Formula	Mass error (ppm)
78.95731	PO_3	-8.19
96.96791	H_2PO_4	-6.31
300.04883	$C_8H_{15}O_9NP$	3.12
845.65894	$C_{58}H_{88}ONP$	-1.02

1005 Fragment ions originating from the precursor molecular ion of 1048 were subjected to
1006 MS/MS fragmentation as described in materials and methods. The proposed molecular
1007 formulae of the product ions are supported by the observed isotope patterns for the
1008 relevant ions.

1009

1010

1011 Table 2. GlcNAc-lipid synthesis in membrane fractions from various bacterial strains.[a]

1012

Enzyme Source	GlcNAc-P-Und, pmol/mg	GlcNAc-P-P-Und, pmol/mg
MGAS5005	259.4±8.2	7.8±1.8
5005Δ*gacI*	Not Detected	15.9±2.5
GBS COH1	62±0.4	5.6±0.8
GBSΔ*gacI*	Not Detected	12±2
E. coli Rosetta (DE3)	Not Detected	19.5±0.3
E. coli:GacI	102.8±3.3	5.4±0.05
E. coli:EpaI	189±3.9	7±1.5

1013

1014 [a]Reaction mixtures contained 50 mM Tris-HCl, pH 7.4, 5 mM 2-mercaptoethanol, 20
1015 mM $MgCl_2$, 1 mM ATP, 5 μM UDP-[^{3}H]GlcNAc (486 cpm/pmol) and bacterial membrane
1016 suspension (50-100 μg membrane protein) in a total volume of 0.01 ml. Following a 10
1017 minute pre-incubation at 30°C, GlcNAc-lipid synthesis was initiated by the addition of
1018 UDP-[^{3}H]GlcNAc. After 5 min, reactions were processed for GlcNAc-lipid synthesis as
1019 described in Materials and Methods. The results are representative of at least three
1020 separate experiments.

1021

1022 Table 3. Kinetic parameters of GlcNAc-lipid synthases in membrane fractions from
1023 various bacterial strains.[a]
1024

Enzyme source	GlcNAc-P-Und		GlcNAc-P-P-Und	
	Km, μM	Vmax, pmol/min/mg	Km, μM	Vmax, pmol/min/mg
MGAS5005	6.4	333	ND	ND
5005Δ*gacI*	ND	ND	19.3	2.75
E. coli:GacI	18.7	54.2	ND	ND
E. coli:GacI/J	1.1	18,500	ND	ND

1025 [a]Membrane fractions from the indicated bacterial strains were assayed for the formation
1026 of [^{3}H]GlcNAc-lipids in the presence of increasing amounts of Und-P, added as a
1027 dispersion in 1% CHAPS. Apparent kinetic parameters for Und-P were calculated from
1028 Michaelis-Menton plots (Fig. S5) using the linear regression algorithm in Sigma Plot ver.
1029 12 (Systat Software, Inc. San Jose, CA). R squared values were 0.9953 for GAS GacI
1030 expressed in *E. coli*; 0.983 for GAS GacI co-expressed with GacJ in *E. coli*, 0.9976 for
1031 GacI in MGAS5005 and 0.931 for GacO in MGAS5005. The results are representative
1032 of at least two separate experiments. ND = Not Detected.
1033

Figure legends:

Figure 1. Map of *S. pyogenes* genes involved in Group A Carbohydrate (GAC) biosynthesis and analysis of 5005Δ*gacI* and 5005 Δ*gacL* deletion mutants. (A) GAC biosynthesis gene cluster. Numbers below represent MGAS5005 gene designations (B) Representative immunoblot analysis of cell-wall fractions isolated from MGAS5005, 5005Δ*gacI*, 5005Δ*gacL* and 5005Δ*gacL gacL*$^{+}$. Data are representative of biological triplicates. (C) Binding of N-acetyl glucosamine specific fluorescein-succinylated-WGA to whole MGAS5005 and 5005Δ*gacL* was measured. Data are the average of three replicates ± standard deviation. (D) Rhamnose and GlcNAc mole percentage of total carbohydrate was determined by GC-MS for cell wall material isolated from MGAS5005 and 5005Δ*gacL* following methanolysis as described in Materials and Methods. Data are the average of four replicates ± standard deviation. The asterisk indicate statistically different values (* $p < 0.05$, ** $p < 0.01$) as determined by the Student's t-test. (E and F) GC-MS chromatograms for glycosyl composition analysis of cell wall isolated MGAS5005 and 5005Δ*gacL*. The deduced schematic structure of the repeating unit of GAC is shown for each strain. The chromatograms are representative of four separate analyses performed on two different cell wall preparations.

Figure 2. Purification and identification of GlcNAc-phosphate-undecaprenol in 5005Δ*gacL.* (A) Thin layer chromatography (TLC) analysis of phospholipids isolated from MGAS5005 and 5005Δ*gacL*. Phospholipids extracted from bacterial strains were separated by TLC on Silica Gel G in $CHCl_3/CH_3OH/H_2O/NH_4OH$ (65:25:4:1). Position of the novel, alkaline-resistant and acid-labile, phospholipid accumulating in 5005Δ*gacL* is indicated by the arrow. The results are representative of three separate experiments.

(B) ESI-MS/MS analysis of the novel phospholipid isolated from 5005Δ*gacL*. The spectrum is assigned to GlcNAc-phosphate-undecaprenol.

Figure 3. Thin layer chromatography of [^{3}H]GlcNAc-lipids from *in vitro* incubations of GAS mutants and *Bacillus cereus* membranes with UDP-[^{3}H]GlcNAc. Membrane fractions from MGAS5005 (panel A), *B. cereus* (Panel B), 5005Δ*gacI* (Panel C), or 5005Δ*gacL* (Panel D) were incubated with UDP-[^{3}H]GlcNAc and analyzed for [^{3}H]GlcNAc lipid synthesis by thin layer chromatography. Reaction mixtures contained 50 mM Tris-Cl, pH 7.4, 5 mM 2-mercaptoethanol, 20 mM $MgCl_2$, 1 mM ATP, 5 μM UDP-[^{3}H]GlcNAc (486 cpm/pmol) and bacterial membrane suspension (100-200 μg membrane protein) in a total volume of 0.02 ml. Following a 10 minute pre-incubation at 30°C, GlcNAc-lipid synthesis was initiated by the addition of UDP-[^{3}H]GlcNAc. After 10 min, reactions were processed for GlcNAc-lipid synthesis as described in Materials and Methods. The organic layers were dried, dissolved in a small volume of $CHCl_3/CH_3OH$ (2:1) and a portion was removed and assayed for radioactivity by liquid scintillation spectrometry. The remainder was spotted on 10 x 20 cm plate of silica gel G and developed in $CHCl_3/CH_3OH/NH_4OH/H_2O$ (65:25:1:4). [^{3}H]GlcNAc-lipids were detected by scanning with an AR2000 Bioscan Radiochromatoscanner. The results are representative of three separate experiments.

Figure 4. Thin layer chromatography of [^{3}H]GlcNAc-lipids from *in vitro* incubations of GAS and GBS membranes with UDP-[^{3}H]GlcNAc. Membrane fractions from WT MGAS5005 (panel A) or GBS COH1 (Panel B) were incubated with UDP-[^{3}H]GlcNAc and analyzed for [^{3}H]GlcNAc-lipid synthesis by thin layer chromatography. Reaction mixtures were exactly as described in the legend to Figure 3. After 5 min incubation with

UDP-[^{3}H]GlcNAc, reactions were processed for GlcNAc-lipid synthesis as described in Materials and Methods. The organic layers were dried, dissolved in a small volume of $CHCl_3/CH_3OH$ (2:1) and a portion was removed and assayed for radioactivity by liquid scintillation spectrometry. The remainder was spotted on 10 x 20 cm plate of silica gel G and developed in $CHCl_3/CH_3OH/NH_4OH/H_2O$ (65:25:1:4). [^{3}H]GlcNAc-lipids were detected by scanning with an AR2000 Bioscan Radiochromatoscanner. The results are representative of three separate experiments.

Figure 5. Effect of ATP on GlcNAc-P-Und synthesis *in vitro* in MGAS5005 membrane fractions. GlcNAc-P-Und synthesis in MGAS5005 membranes was assayed after the indicated time at 30 °C in the presence (○) or absence (●) of 1 mM ATP. Reaction mixtures were identical to those described in Figure 3 except for the presence of ATP. Following incubation, incorporation into [^{3}H]GlcNAc-P-Und was determined as described in Materials and Methods. The results are representative of three separate

Figure 6. GacI and GacJ exist as a detergent-stable complex in the membrane. GacJ and His-tagged GacI were co-expressed in *E. coli* Rosetta DE3 cells and extracted from the membrane fraction in 2.5% CHAPS. The proteins were purified using Ni-NTA agarose in the presence of 2.5% CHAPS. (A) Fractions collected during Ni-NTA purification were analyzed by immunoblot using anti-His antibodies. (B) The eluted proteins were analyzed by SDS-PAGE. The results are representative of three separate experiments.

Figure 7. Analysis of GacO function in MGAS5005. (A) GacO catalyzes the synthesis of GlcNAc-P-P-Und in *E. coli* membranes**.** Reaction mixtures contained 50 mM Tris-HCl, pH 7.4, 5 mM 2-mercaptoethanol, 20 mM $MgCl_2$, 0.5 % CHAPS, 20 μM Und-P

(dispersed by ultrasonication in 1 % CHAPS), 5 μM UDP-GlcNAc (452 cpm/pmol) and *E. coli* membrane fraction from either PR4019, CLM37 or CLM37:GacO strains. Following incubation for 10 min at 30 ºC, incorporation of [^{3}H]GlcNAc into [^{3}H]GlcNAc-P-P-Und was determined as described in Materials and Methods. Data are the average of three replicates ± standard deviation.

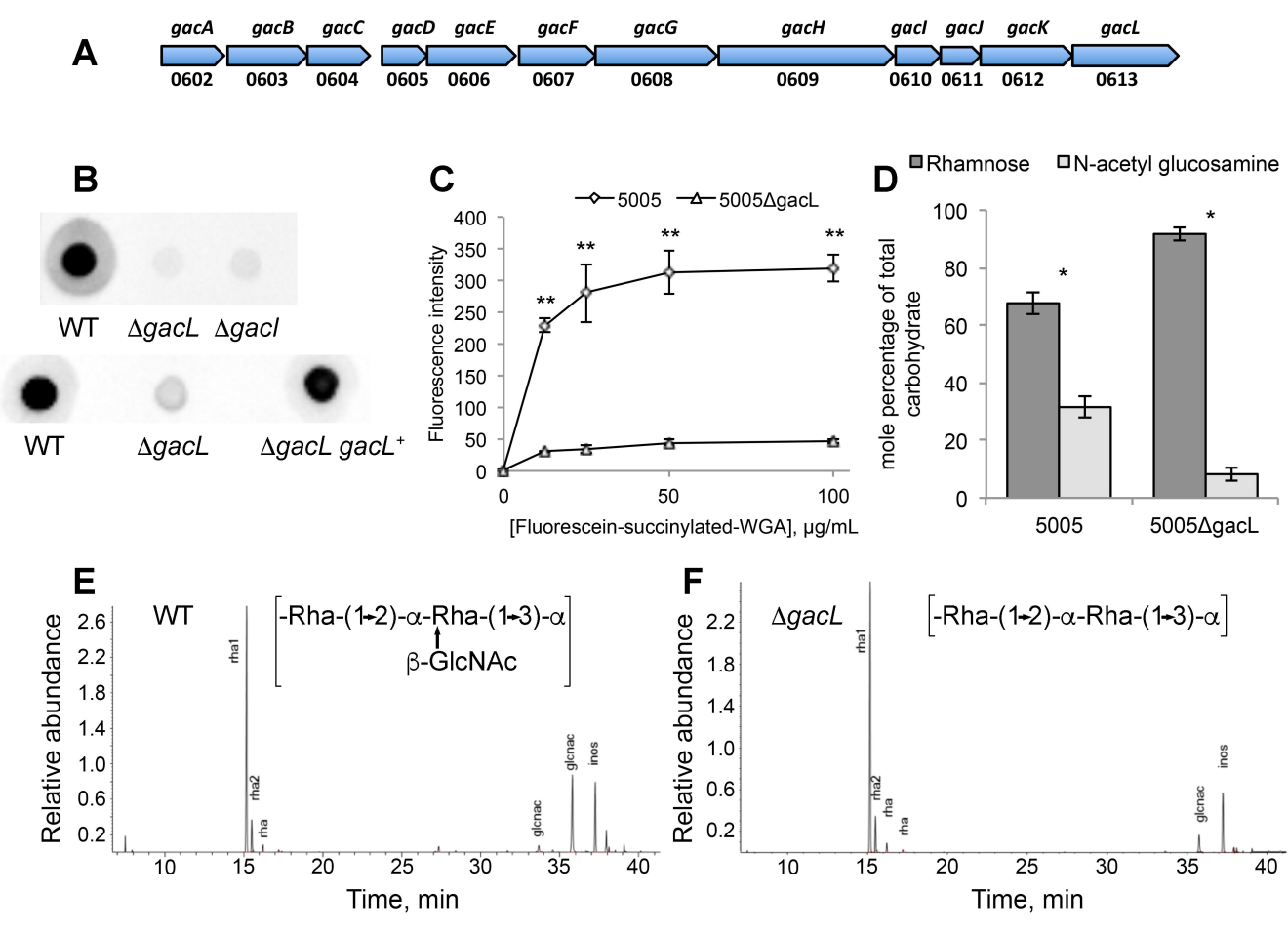
(B and C) GlcNAc-P-P-Und can function as an acceptor substrate for rhamnosylation to form Rhamnosyl-GlcNAc-P-P-Und in 5005Δ*gacI* membranes**.** Incubation conditions were as described in the legend to Figure 3. (B) After 35 minutes incubation with UDP-[^{3}H]GlcNAc, the reactions were analyzed for formation of [^{3}H]GlcNAc lipids as described in Figure 3. (C) After 5 minutes incubation with UDP-[^{3}H]GlcNAc, the reactions were incubated with 20 μM TDP-Rhamnose for an additional 30 min and analyzed for formation of [^{3}H]GlcNAc lipids as described in Figure 3. The results are representative of three separate experiments.

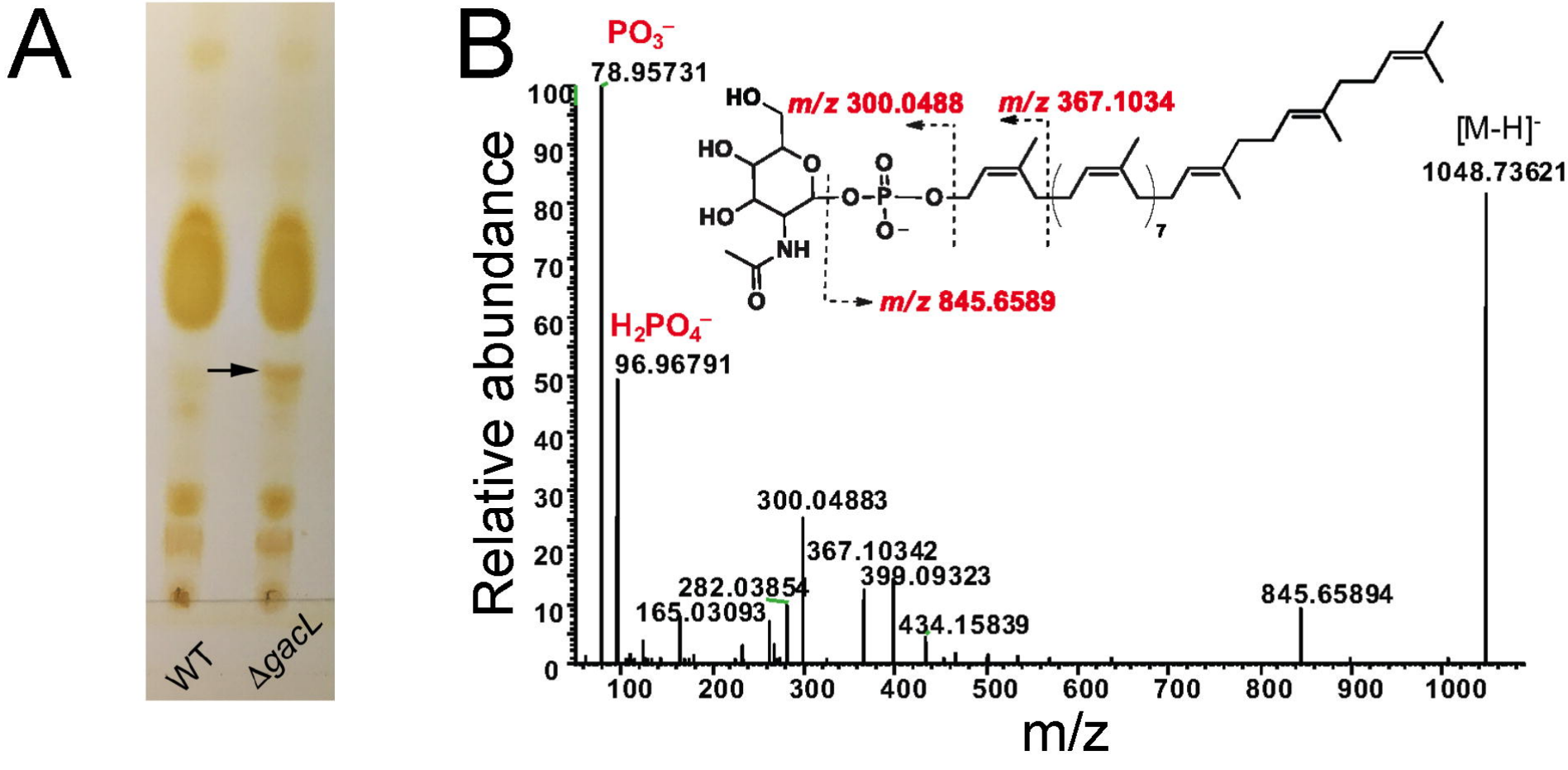
Figure 8. The absence of GacL increases sensitivity to cell wall amidases. Mid-exponential phase MGAS5005 and 5005Δ*gacL* were grown in the indicated concentrations of (A) CbpD, (B) PlyPy, and (C) PlyC. The change in growth is represented as a percentage of growth where no amidase was present. Data are the average of three replicates ± standard deviation. The asterisk indicate statistically different values (** $p < 0.01$) as determined by the Student's t-test.

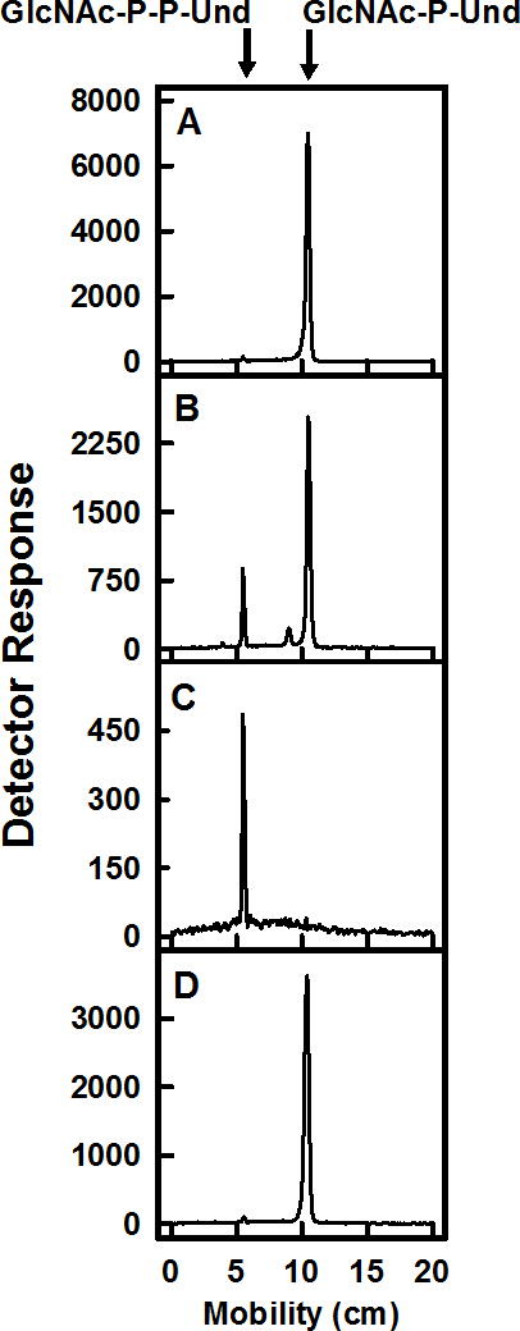
Figure 9. Schematic diagram of GAC biosynthesis. GAC is anchored to peptidoglycan presumably via phosphodiester bond. GAC biosynthesis is initiated on the inner leaflet of the plasma membrane where GacO produces GlcNAc-P-P-Und which serves as a membrane-anchored acceptor for polyrhamnose synthesis catalyzed by the GacB,

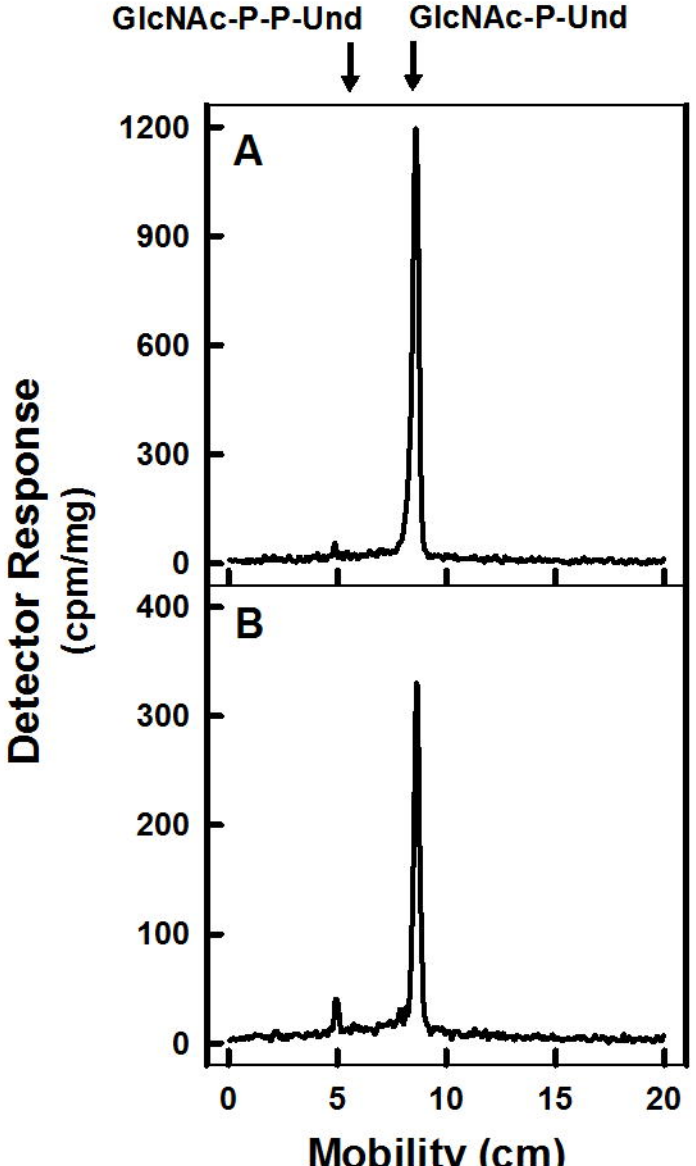
GacC, GacF and GacG rhamnosyltransferases. Following polymerization polyrhamnose is transferred to the outer leaflet of the membrane presumably by the GacD/ GacE ABC transporter. Also in the inner leaflet of the membrane, GacI aided by GacJ produces GlcNAc-P-Und which then diffuses across the plasma membrane to the outer leaflet aided by GacK. Subsequently, GacL transfers GlcNAc to polyrhamnose using GlcNAc-P-Und as glycosyl donor. Lastly, protein members of LytR-CpsA-Psr phosphotransferase family presumably attach GAC to peptidoglycan. Several details of this biosynthetic scheme are still speculative and further research will be required to definitively confirm this hypothetical pathway, but the overall organization is consistent with other isoprenol-mediated capsular polysaccharide pathways.

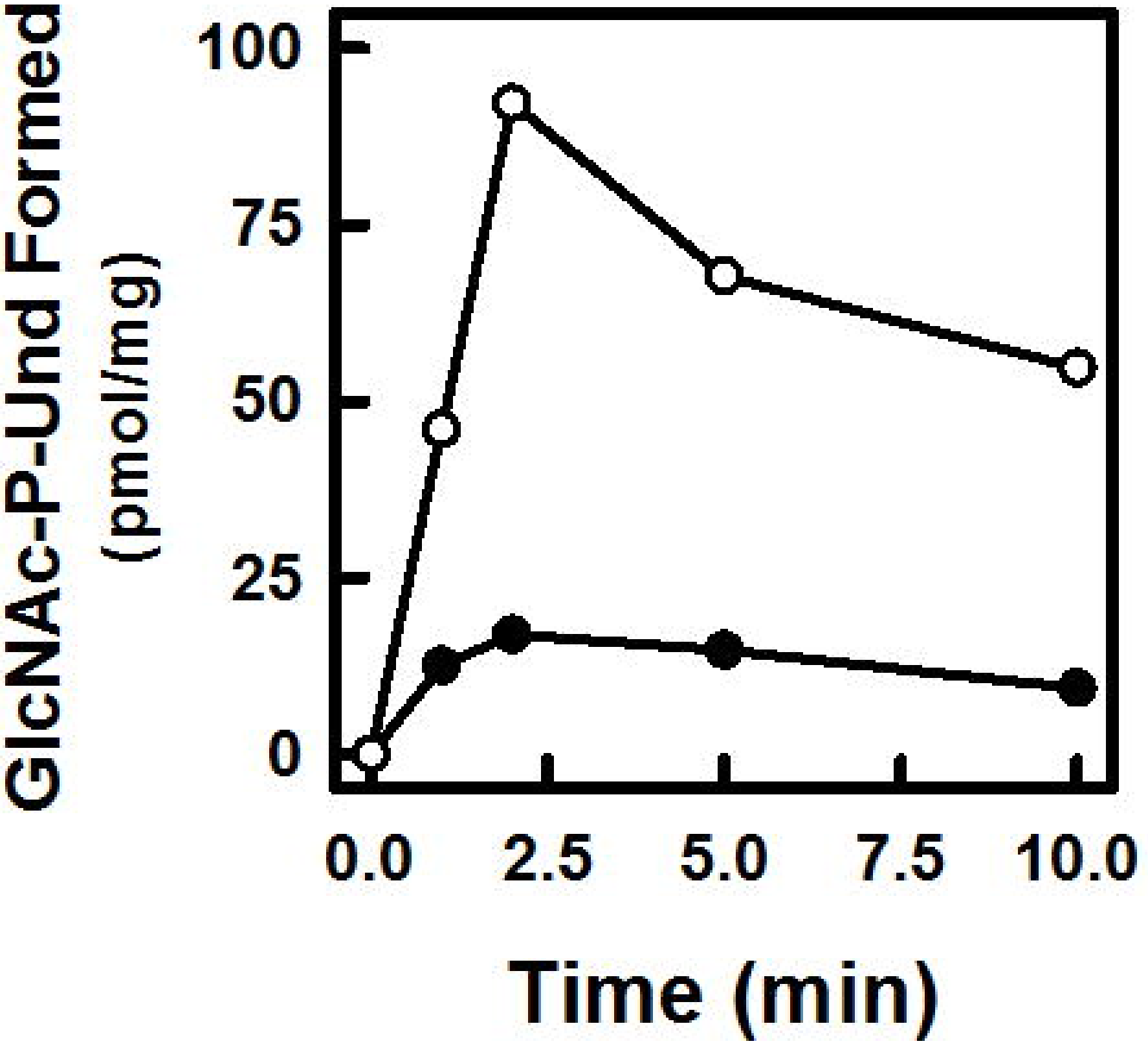
Figure 10. Sequence relationship of GacI family of proteins. Homology between GacI homologs is graphically displayed using CLANS analysis (40). Dots correspond to individual protein sequences selected as described in Material and Methods and provided as Supplementary File. Selected homologs of *S. pyogenes* are highlighted by colored dots.

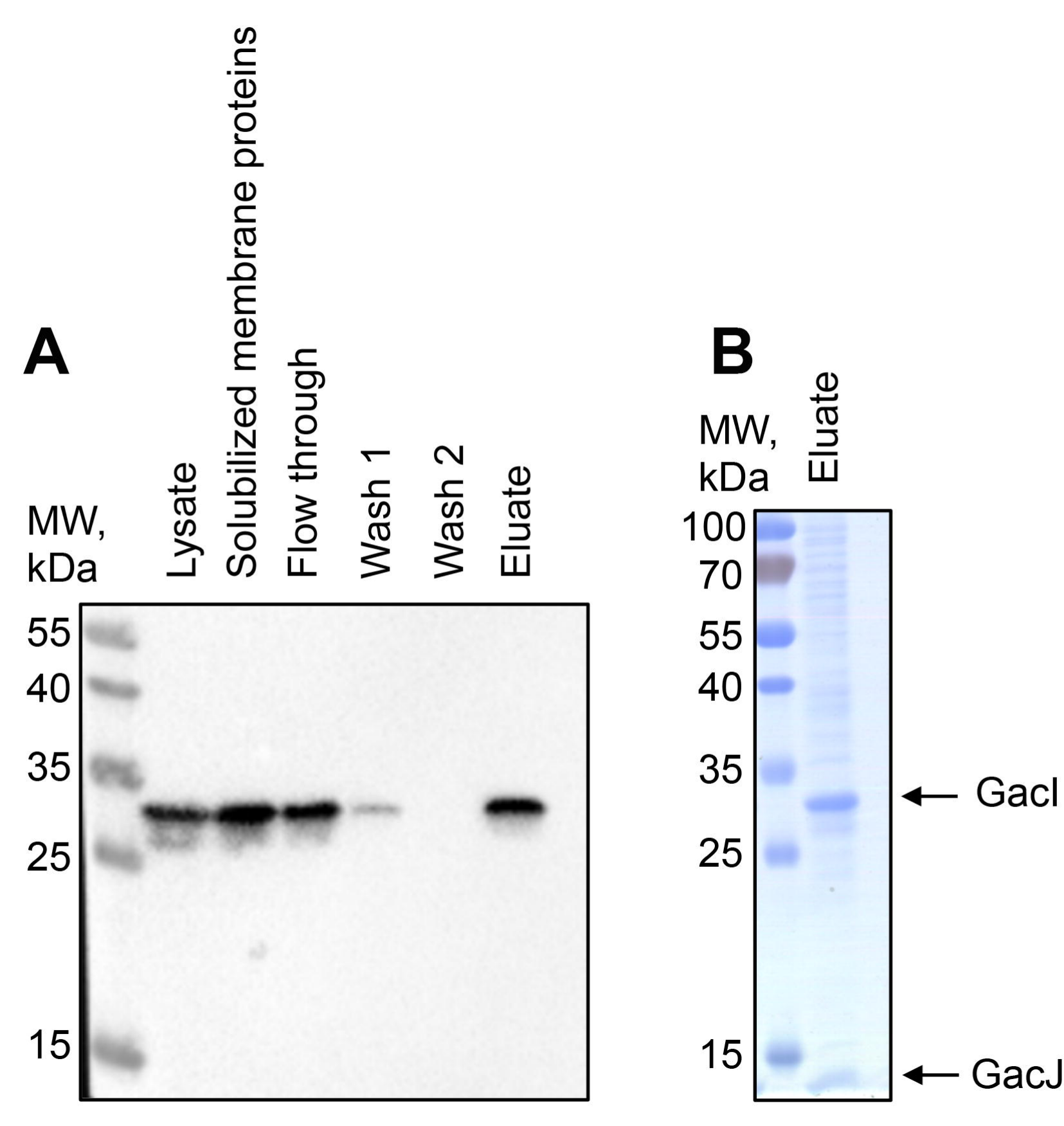

A
gacA gacB gacC gacD gacE gacF gacG gacH gacI gacJ gacK gacL
0602 0603 0604 0605 0606 0607 0608 0609 0610 0611 0612 0613
B
WT ΔgacL ΔgacI
WT ΔgacL ΔgacL gacL+
C
5005 5005ΔgacL
Fluorescence intensity
[Fluorescein-succinylated-WGA], μg/mL
D
Rhamnose N-acetyl glucosamine
mole percentage of total carbohydrate
5005 5005ΔgacL
E
WT
-Rha-(1-2)-α-Rha-(1-3)-α
β-GlcNAc
Relative abundance
Time, min
F
ΔgacL
-Rha-(1-2)-α-Rha-(1-3)-α
Relative abundance
Time, min

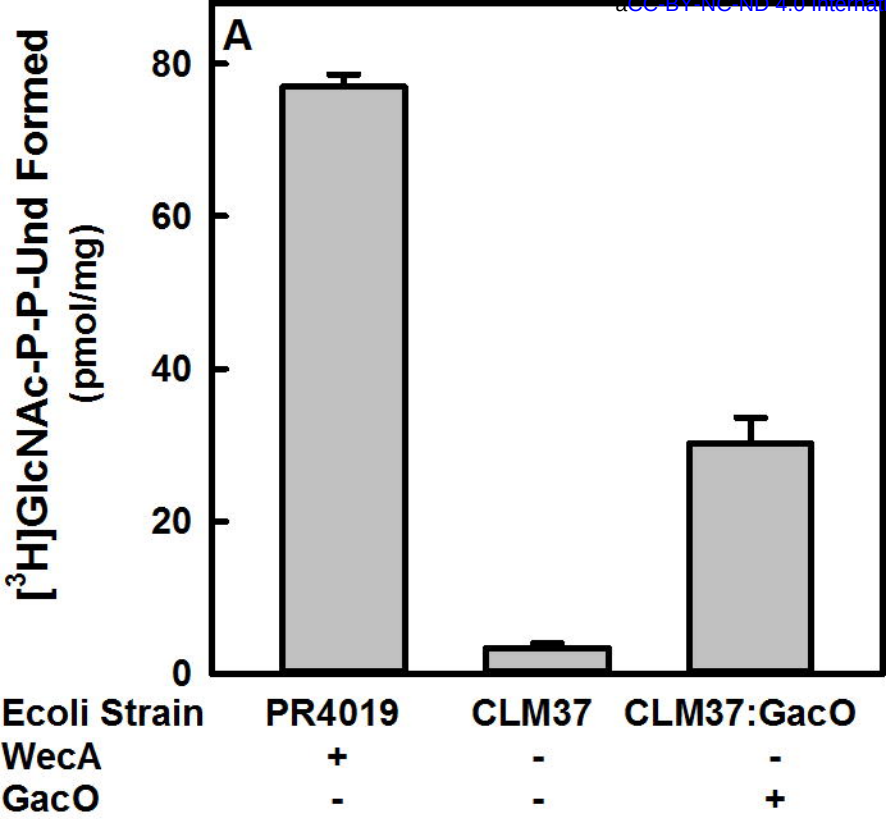

A
WT
ΔgacL
B
PO3−
78.95731
m/z 300.0488
m/z 367.1034
[M-H]-
1048.73621
m/z 845.6589
H2PO4−
96.96791
Relative abundance
300.04883
367.10342
399.09323
282.03854
165.03093
434.15839
845.65894
m/z

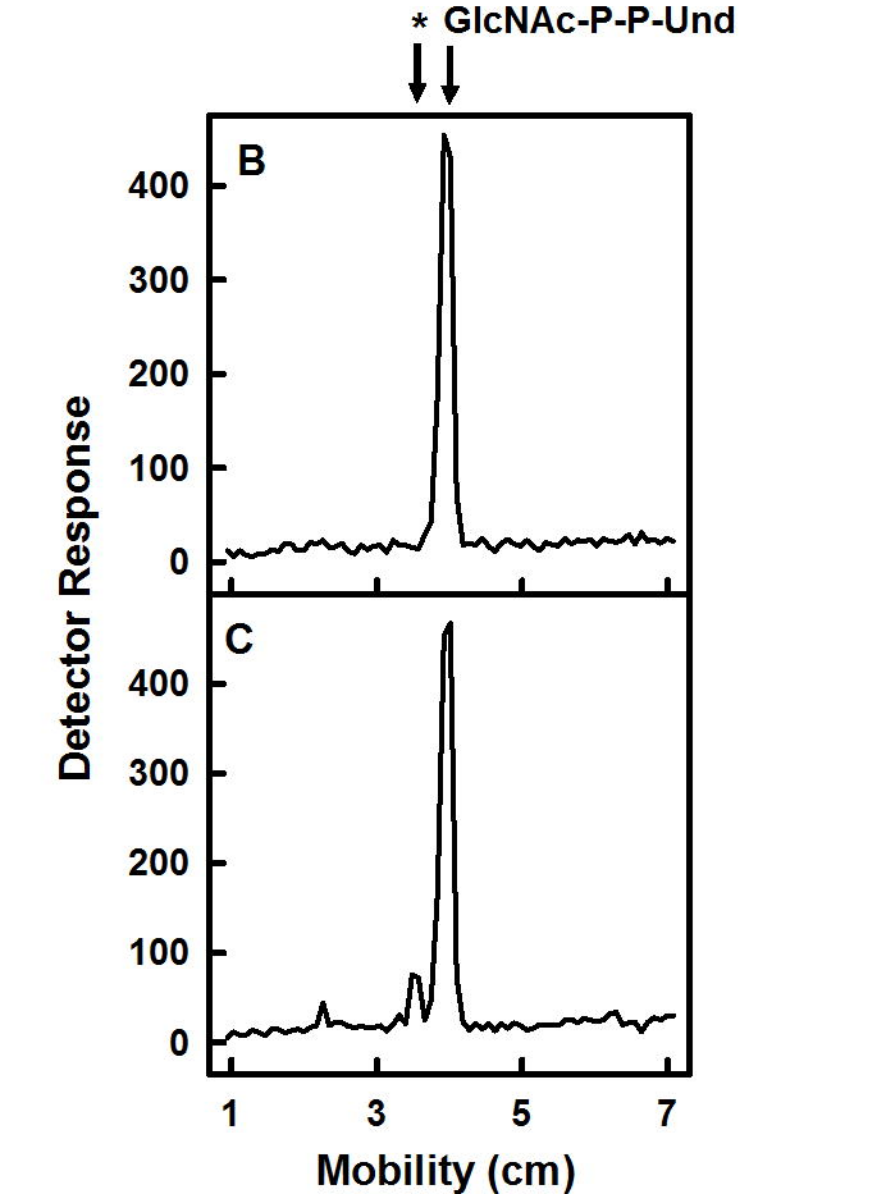

GlcNAc-P-P-Und
GlcNAc-P-Und
A
B
C
D
8000
6000
4000
2000
0
2250
1500
750
0
450
300
150
0
3000
2000
1000
0
Detector Response
0
5
10
15
20
Mobility (cm)

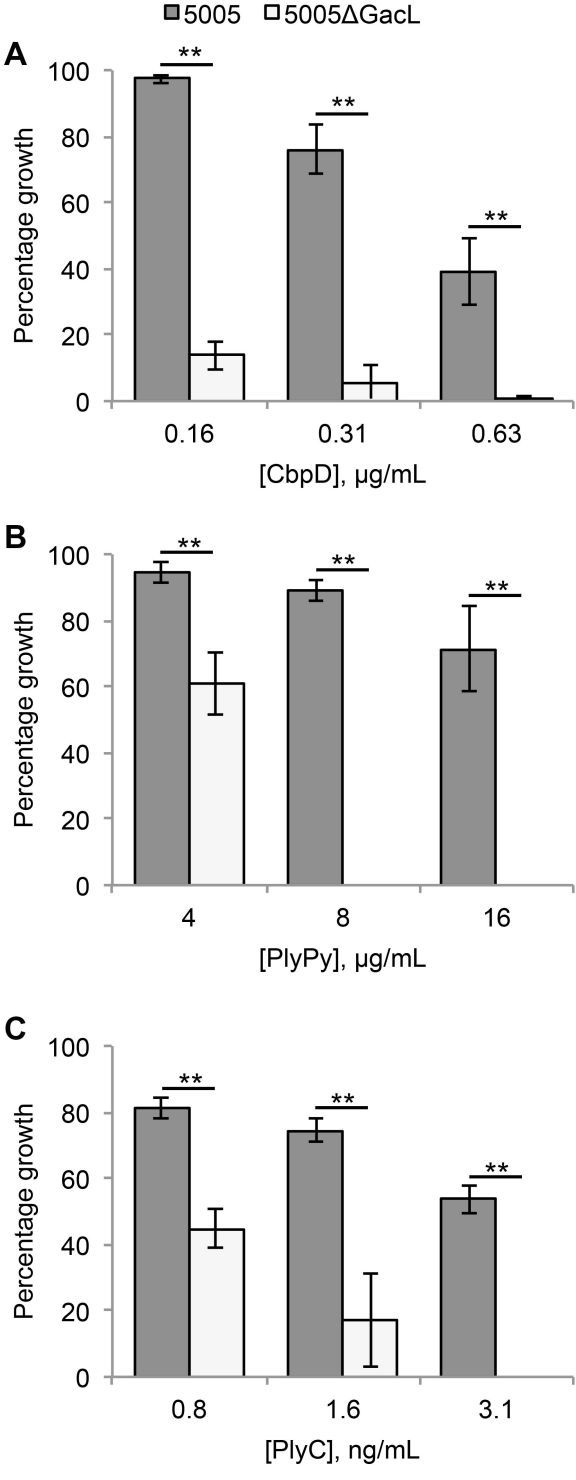

GlcNAc-P-P-Und
GlcNAc-P-Und
A
B
1200
900
600
300
0
400
300
200
100
0
0
5
10
15
20
Detector Response
(cpm/mg)
Mobility (cm)

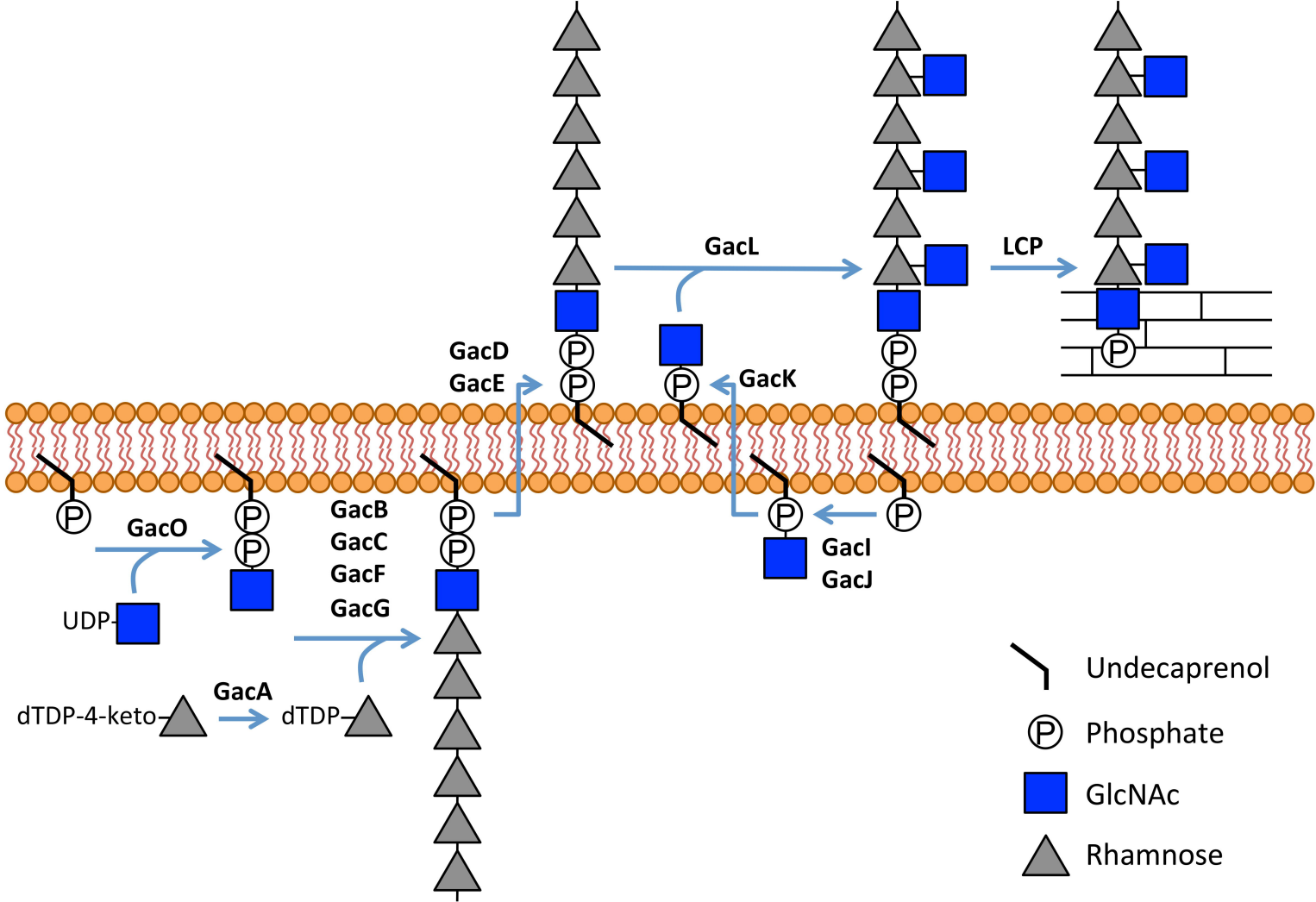

GlcNAc-P-Und Formed
(pmol/mg)
100
75
50
25
0
0.0
2.5
5.0
7.5
10.0
Time (min)

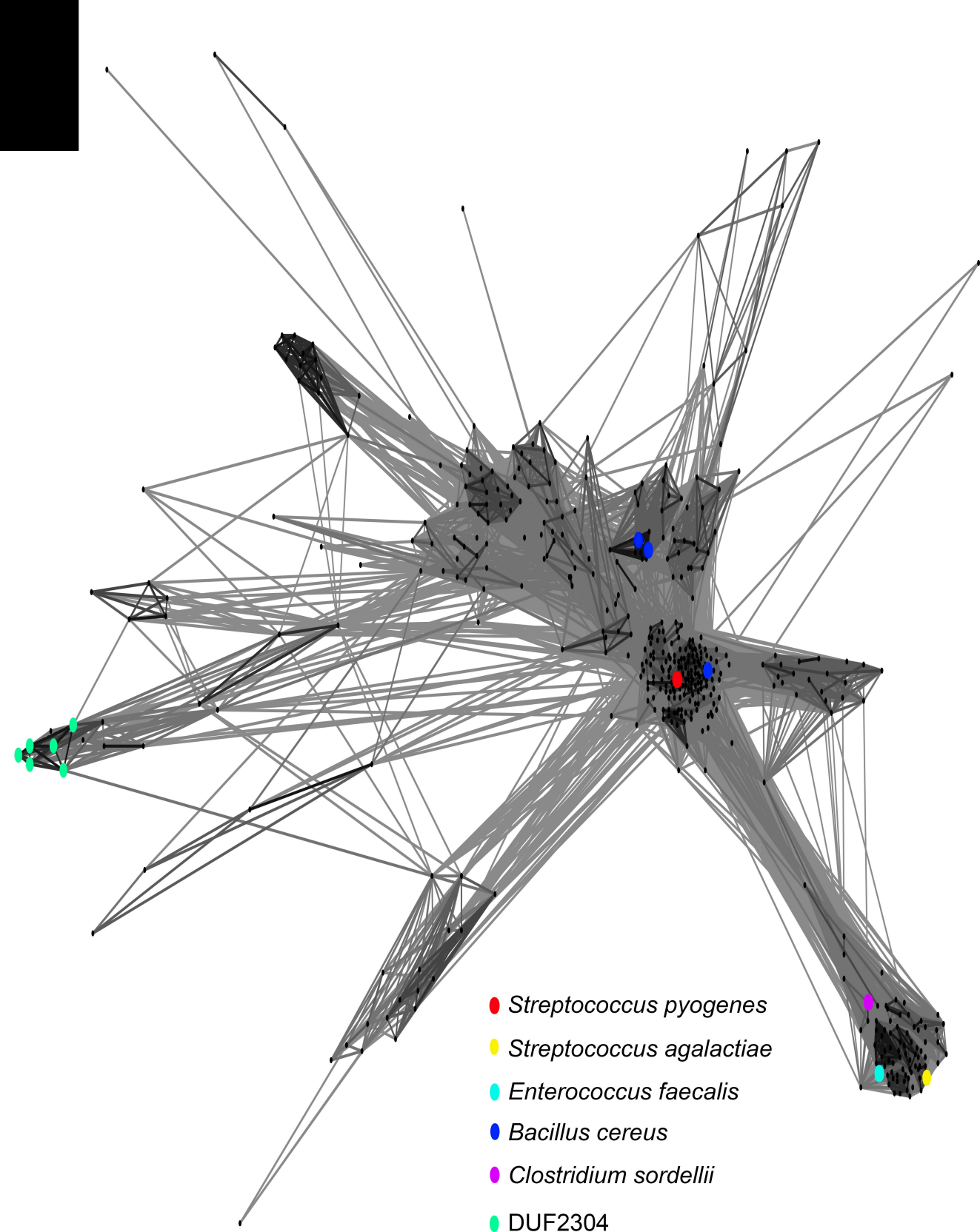

A
MW, kDa
Lysate
Solubilized membrane proteins
Flow through
Wash 1
Wash 2
Eluate
55
40
35
25
15
B
MW, kDa
Eluate
100
70
55
40
35
25
15
GacI
GacJ






5005
5005ΔGacL
A
Percentage growth
100
80
60
40
20
0
**
**
**
0.16
0.31
0.63
[CbpD], µg/mL
B
Percentage growth
100
80
60
40
20
0
**
**
**
4
8
16
[PlyPy], µg/mL
C
Percentage growth
100
80
60
40
20
0
**
**
**
0.8
1.6
3.1
[PlyC], ng/mL


GacL
LCP
GacD
GacE
GacK
GacO
GacB
GacC
GacF
GacG
GacI
GacJ
UDP-
GacA
dTDP-4-keto-
dTDP-
Undecaprenol
Phosphate
GlcNAc
Rhamnose


Streptococcus pyogenes
Streptococcus agalactiae
Enterococcus faecalis
Bacillus cereus
Clostridium sordellii
DUF2304