nature chemical biology

ARTICLES

https://doi.org/10.1038/s41589-019-0251-4

Discovery of glycerol phosphate modification on streptococcal rhamnose polysaccharides

Rebecca J. Edgar[1,9], Vincent P. van Hensbergen[2,11], Alessandro Ruda[3,11], Andrew G. Turner[4], Pan Deng[5], Yoann Le Breton[6,10], Najib M. El-Sayed[6,7], Ashton T. Belew[6,7], Kevin S. McIver[6], Alastair G. McEwan[4], Andrew J. Morris[5], Gérard Lambeau[8], Mark J. Walker[4], Jeffrey S. Rush[1], Konstantin V. Korotkov[1], Göran Widmalm[3], Nina M. van Sorge[2,12*] and Natalia Korotkova[1,12*]

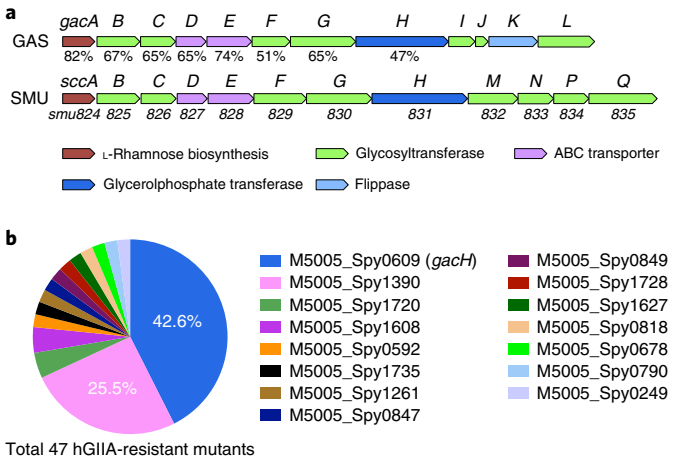
Cell wall glycopolymers on the surface of Gram-positive bacteria are fundamental to bacterial physiology and infection biology. Here we identify *gacH*, a gene in the *Streptococcus pyogenes* group A carbohydrate (GAC) biosynthetic cluster, in two independent transposon library screens for its ability to confer resistance to zinc and susceptibility to the bactericidal enzyme human group IIA-secreted phospholipase A_2. Subsequent structural and phylogenetic analysis of the GacH extracellular domain revealed that GacH represents an alternative class of glycerol phosphate transferase. We detected the presence of glycerol phosphate in the GAC, as well as the serotype *c* carbohydrate from *Streptococcus mutans*, which depended on the presence of the respective *gacH* homologs. Finally, nuclear magnetic resonance analysis of GAC confirmed that glycerol phosphate is attached to approximately 25% of the GAC *N*-acetylglucosamine side-chains at the C6 hydroxyl group. This previously unrecognized structural modification impacts host–pathogen interaction and has implications for vaccine design.

Gram-positive bacteria are surrounded by a thick cell wall consisting of a complex network of peptidoglycan with covalently attached glycopolymers that comprise a large family of structurally diverse molecules, including wall teichoic acid, mycobacterial arabinogalactans and capsular polysaccharides. From these, wall teichoic acid is perhaps the most widespread and best-studied molecule. This polyanionic, phosphate-rich glycopolymer is critical for functions including cell division, antibiotic resistance, metal ion homeostasis, phage-mediated horizontal gene transfer and protection of bacteria from host defense peptides and antimicrobial enzymes[1,2]. As such, these structures and their biosynthetic pathways are attractive targets for antibiotic development and vaccine design. Interestingly, many streptococci lack classical wall teichoic acid and instead express glycopolymers that are characterized by the presence of L-rhamnose (Rha)[3]. These structures comprise about 40–60% of the bacterial cell wall mass, and are historically used for serological grouping of streptococci[3]. The glycopolymers of two human streptococcal pathogens, *S. pyogenes* or group A *Streptococcus* (GAS) and *S. mutans*, are respectively referred to as group A carbohydrate (GAC) and serotype *c* carbohydrate (SCC). These glycopolymers share a characteristic (→3)α-Rha(1→2)α-Rha(1→) polyrhamnose backbone, but are serologically distinguished based on their specific glycosyl side-chain residues, that is N-acetyl-β-D-glucosamine (GlcNAc) in GAC[4] and α-glucose (Glc) in SCC[5]. GAC and SCC play key roles in cell morphology and division[6], resistance to certain cell wall-targeting antibiotics[7], biofilm formation[8] and pathogenesis of GAS and *S. mutans*[9–11]. Importantly for both pathogens, GAC and SCC have been evaluated as vaccine antigens. Immunization with GAC or SCC induces opsonophagocytic antibodies that enhance killing of GAS and *S. mutans*, respectively[5,12,13]. In addition, GAC has proven efficacious as a vaccine antigen through active immunization in mice[12,13].

The GAC and SCC biosynthetic pathways are encoded by 12-gene clusters[3,11], herein designated as *gacABCDEFGHIJKL* and *sccABCDEFGHMNPQ* (Fig. 1a), respectively. The first seven genes in both clusters are conserved in many streptococcal species, and they participate in polyrhamnose backbone synthesis and transport[14]. In GAS, *gacI*, *gacJ*, *gacK* and *gacL* encode the machinery to generate and add the GlcNAc side-chain to the polyrhamnose backbone[11,15], whereas the genes required for Glc side-chain generation are not clearly identified in *S. mutans*. In addition to these streptococcal species, similar gene clusters are present in a wide variety of streptococcal, lactococcal and enterococcal species[3].

In addition to the polyrhamnose biosynthesis genes, the GAC and SCC biosynthetic clusters contain further conserved genes of unknown function, *gacH* and *sccH*, respectively, which are annotated as putative glycerol phosphate (GroP) transferases. Recently, we employed the *Krmit* GAS transposon mutant library[16] and identified *gacI* and *gacH* as genes that confer bacterial sensitivity to human group IIA-secreted phospholipase A2 (hGIIA)[17], an important bactericidal protein of the innate immune system against Gram-positive pathogens[18]. Complementary to that study, we now

[1]Department of Molecular and Cellular Biochemistry, University of Kentucky, Lexington, KY, USA. [2]Department of Medical Microbiology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands. [3]Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, Stockholm, Sweden. [4]Australian Infectious Diseases Research Centre and School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Australia. [5]Division of Cardiovascular Medicine and the Gill Heart Institute, University of Kentucky, Lexington, KY, USA. [6]Department of Cell Biology & Molecular Genetics and Maryland Pathogen Research Institute, University of Maryland, College Park, MD, USA. [7]Center for Bioinformatics and Computational Biology, University of Maryland, College Park, MD, USA. [8]Université Côte d'Azur, Centre National de la Recherche Scientifique, Institut de Pharmacologie Moléculaire et Cellulaire, Valbonne Sophia Antipolis, France. [9]Present address: Callaghan Innovation, Gracefield, Lower Hutt, New Zealand. [10]Present address: Wound Infections Department, Bacterial Diseases Branch, Walter Reed Army Institute of Research, Silver Spring, MD, USA. [11]These authors contributed equally: Vincent P. van Hensbergen, Alessandro Ruda. [12]These authors jointly supervised this work: Nina M. van Sorge and Natalia Korotkova. *e-mail: nsorge3@umcutrecht.nl; nkorotkova@uky.edu



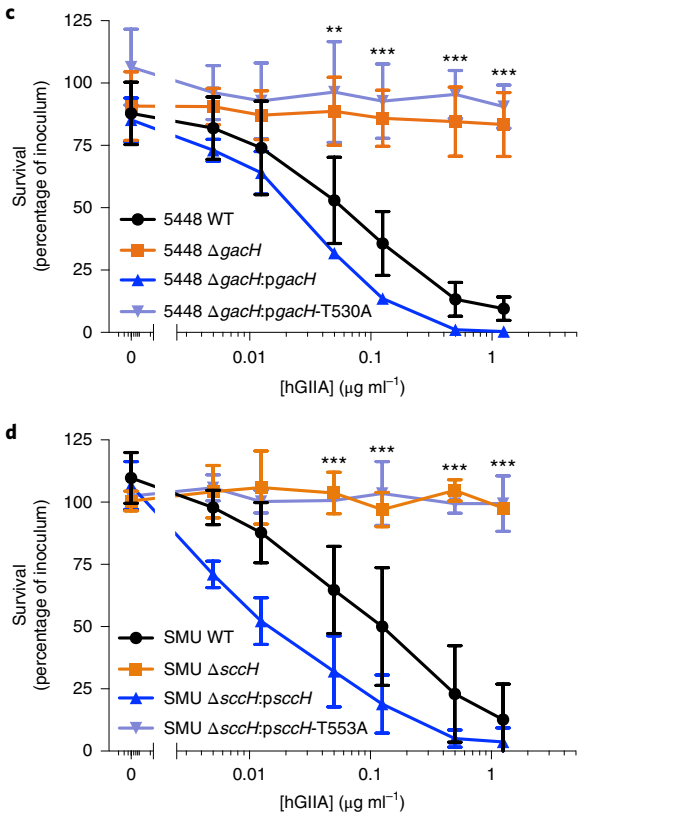


Fig. 1 | GacH homologs are required for hGIIA bactericidal activity against GAS and *S. mutans*. a, Schematic representation of GAC and SCC biosynthetic gene clusters. SCC biosynthesis encoding gene cluster *smu*.824–835 was renamed *sccABCDEFGHMNPQ*. Sequence identity (%) between encoded homologous proteins is indicated. Sequences of GAS 5005 and *S. mutans* UA159 were used for identity comparison. **b**–**d**, Identification of *gacH* in transposon sequencing (Tn-seq) screen and validation for hGIIA resistance. Transposon gene locus tags of the 47 hGIIA-resistant mutants after exposure of *Krmit* mutant transposon library to lethal concentrations of hGIIA (**b**). Susceptibility of GAS 5448 and *S. mutans* to hGIIA concentration range following deletion of *gacH* in GAS 5448 (**c**) and the *gacH*-homologous gene *sccH* (**d**), respectively. Symbols and error bars represent the mean and s.d., respectively ($n = 3$ biologically independent replicates and each replicate represents three technical replicates). *P* values were calculated by two-way analysis of variance (ANOVA). Bonferroni multiple comparison test was used to statistically compare multiple groups. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$. The precise *P* values are listed in Supplementary Table 2.

identify *gacH* as the only valid hit when the *Krmit* library was exposed to a lethal concentration of hGIIA. Interestingly, *gacH* was also identified as a gene providing resistance to zinc toxicity, which is a mechanism deployed by neutrophils to kill GAS[19]. In pursuit of the underlying mechanism, we have characterized the function of GacH at the genetic, biochemical and structural level. Our study identifies a previously overlooked GroP modification on both GAC and SCC, and pinpoints GacH homologs as the enzymes responsible for the respective GroP modifications.

Results

GacH and SccH confer sensitivity to hGIIA. We previously identified *gacH* in a GAS Tn-seq transposon library screen as a potential hGIIA susceptibility gene[17]. To identify additional resistant mutants, we exposed the *Krmit* GAS transposon library[16] to a lethal concentration of hGIIA. PCR sequencing identified that 43% of the recovered mutants had a transposon insertion in *gacH*, and 26% in *M5005_Spy_1390* (Fig. 1b and Supplementary Table 1). *M5005_Spy_1390* was identified in the initial susceptibility screen as an artifact due to biased transposon insertions[17] and was not investigated further. To validate our finding for *gacH*, we generated a *gacH* deletion mutant in a GAS serotype M1T1 clone 5448, creating 5448Δ*gacH*. Deletion of *gacH* rendered GAS resistant to hGIIA over the tested concentration range, and was reversed by complementation with *gacH* on an expression plasmid (5448Δ*gacH*:p*gacH*, Fig. 1c). This *gacH*-dependent hGIIA resistance was also observed in two different GAS backgrounds, 2221 (M1T1 clone strain) and 5005 (clinical *covS* mutant isolate of M1T1 strain) (Supplementary Fig. 1a,b), demonstrating that the effect is conserved across GAS strains of the M1T1 background and independent of CovRS status—a two-component system that regulates about 15% of the genes in this bacterium[20].

To investigate whether hGIIA susceptibility was also influenced by *gacH* homologs in other streptococci, we deleted *sccH* in *S. mutans* (SMU) serotype *c* strain Xc, creating SMUΔ*sccH*. SMUΔ*sccH* was completely resistant to the tested hGIIA concentrations (Fig. 1d) and susceptibility was restored to wild-type levels by plasmid-expressed *sccH*. However, heterologous expression of *gacH* in SMUΔ*sccH* did not restore the phenotype (Supplementary Fig. 1c), suggesting that the enzymes may target different substrates. Altogether, our data indicate that deletion of *gacH* homologs renders streptococci more resistant to hGIIA bactericidal activity and that GacH function is species-specific.

GacH and SccH provide protection from zinc toxicity. Recent evidence indicates that neutrophils deploy zinc poisoning as an antimicrobial strategy against GAS during phagocytosis[19]. To resist Zn^{2+} toxicity, GAS expresses a zinc efflux system encoded by *czcD*[19]. To search for additional Zn^{2+}-resistance genes, we performed a Tn-seq screen of the GAS *Krmit* transposon library[16] using two Zn^{2+} concentrations, 10 and 20 μM, selected based on growth inhibition analysis (Supplementary Fig. 2a). Genomic DNA for Tn-seq analysis was collected after T_2 and T_3 passages (Supplementary Fig. 2b). In addition to the expected importance of *czcD*, *gacI* and *gacH* transposon insertions were significantly reduced in the library ($P < 0.05$) after growth with 20 μM Zn^{2+} in both T_2 and T_3 passages compared to untreated controls, indicating that these genes provide resistance against Zn^{2+} toxicity (Fig. 2a–d).

To validate our findings, we grew 5448Δ*gacH* and 5448Δ*gacI*[11] on solid rich medium supplied with different Zn^{2+} concentrations (Fig. 2e,f). Both mutants showed reduced growth in the medium supplied with 1.25 mM Zn^{2+}, which was restored following complementation with the respective genes (Fig. 2e,f). Again, we checked for function conservation by extending our experiments to *S. mutans*. Indeed, SMUΔ*sccH* was more sensitive to Zn^{2+} in comparison to the parental strain and the phenotype could be restored by *sccH* but not by *gacH* (Supplementary Fig. 3). Hence, our results

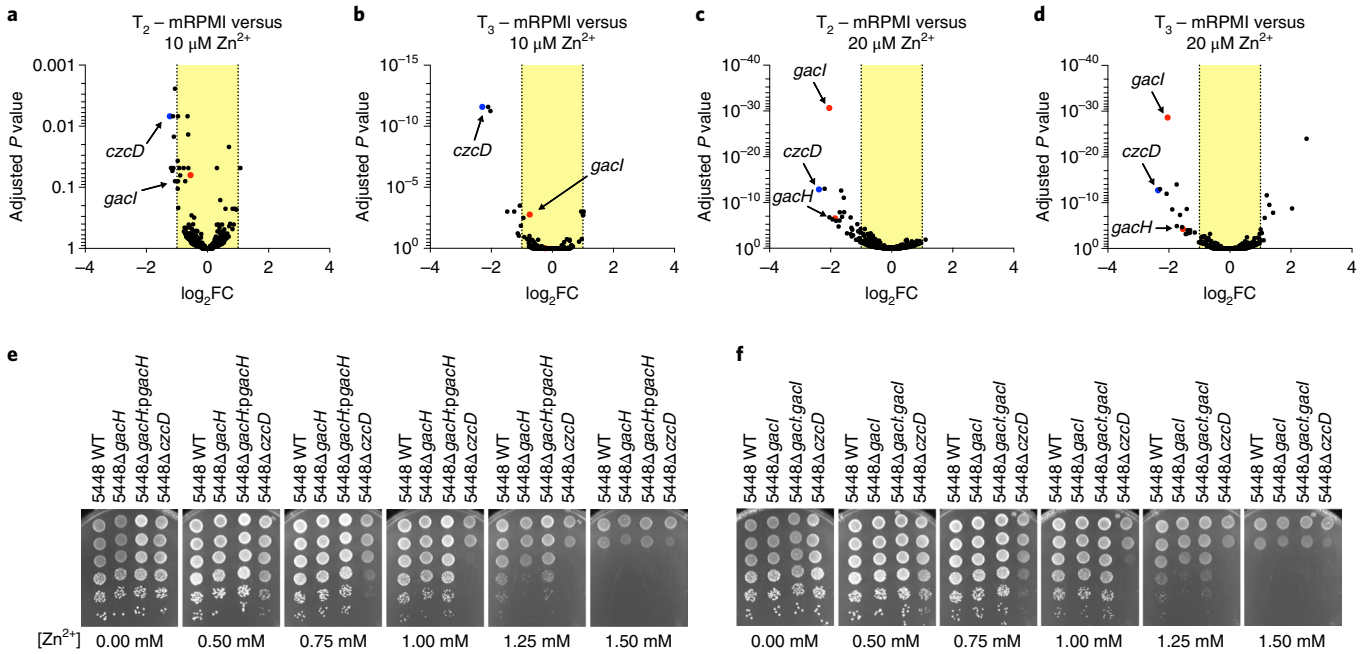


Fig. 2 | Deletion of *gacI* and *gacH* renders GAS susceptible to Zn^{2+}. a–d, Tn-seq volcano plots showing representation of *czcD*, *gacH* and *gacI* in GAS *Krmit* transposon library screens for Zn^{2+} tolerance. log_2 fold change (log_2FC) in fitness was plotted against adjusted *P* values from Tn-seq analyses through an established pipeline using EdgeR and DEseq2 ($n = 4$ biologically independent replicates used for analysis)[47–49]. The outline of the experiment is shown in Supplementary Fig. 2b. Tn-seq screens of the transposon library were conducted using 10 μM Zn^{2+} at T_2(**a**), 10 μM Zn^{2+} at T_3 (**b**), 20 μM Zn^{2+} at T_2 (**c**) and 20 μM Zn^{2+} at T_3(**d**). **e**,**f**, Zn^{2+} sensitivity as tested by drop-test assay using strains GAS 5448 wild type, 5448Δ*gacH* and 5448Δ*gacH*:p*gacH* (**e**); and GAS 5448 wild type, 5448Δ*gacI* and 5448Δ*gacI*:*gacI* (**f**). 5448Δ*czcD* was included as a positive control in both panels. Each drop-test assay experiment was performed independently at least three times and yielded the same results.

provide strong evidence that GacH and SccH are important in protecting streptococci from Zn^{2+} toxicity.

Crystal structure reveals that GacH is a GroP transferase. GacH is predicted to contain 11 N-terminal transmembrane segments and an extracellular C-terminal catalytic domain (eGacH). To test the hypothesis that GacH is a GroP transferase, eGacH was expressed and purified from *Escherichia coli*. Its crystal structure was determined in apo form (Protein Data Bank (PDB) No. 5U9Z) at 2.0 Å resolution (Fig. 3a,b) and in complex with GroP (PDB No. 6DGM) at 1.49 Å resolution (Fig. 3c). The apo- and GroP-containing eGacH structures belong to different crystal forms, with two molecules in the asymmetric unit. Analysis of the dimer interface and other crystal contacts revealed that the dimer interface has the largest surface of all crystal contacts (1809 and 1894 $Å^2$ in the two crystal forms). However, it is scored below the stable complex formation criteria and recombinant eGacH behaves as a monomer in solution. The structures of the apo- and GroP-bound eGacH monomers are very similar, with root mean square deviation of 0.3 Å for 380 superimposed Cα atoms, as well as between the non-crystallographic copies.

The extracellular C-terminal catalytic domain has an α/β core structure that is characteristic for the sulfatase protein family, with the closest similarity to lipoteichoic acid (LTA) synthase LtaS[21,22] (Supplementary Figs. 4a,b and 5) and LTA primase LtaP[23] (Supplementary Table 3). LtaS and LtaP are GroP transferases that participate in biosynthesis of LTA, a crucial constituent of Gram-positive cell envelopes, consisting of a poly(GroP) backbone linked to a glycolipid membrane anchor[24]. The catalytic site of eGacH contains a Mn^{2+} ion coordinated by residues E488, T530, D711 and H712, equivalent to residues E255, T300, D475 and H476 of a C-terminal extracellular domain of LtaS (eLtaS) from *Staphylococcus aureus* (Fig. 3c and Supplementary Figs. 4c,d and 5). The structure of eGacH in complex with GroP revealed the position of the ligand in the active site with the phosphoryl group oriented towards the Mn^{2+} ion, and coordinated by residues G529, T530 and H650 (Fig. 3c). The glycerol 2- and 3-hydroxyl groups form hydrogen bonds with side-chains of residues R589, H580 and N586. The positions of GroP and coordinating residues are similar in eGacH and *S. aureus* eLtaS structures. For example, the glycerol moiety forms hydrogen bonds with residues H580 and R589 in GacH and equivalent residues H347 and R356 in *S. aureus* eLtaS (Fig. 3c and Supplementary Fig. 4c,d)[21]. Thus, the structure of eGacH in complex with GroP is consistent with the idea that GacH and LtaS use related catalytic mechanisms to transfer GroP to substrates.

To functionally assess the requirement of the catalytic residues, we examined the bactericidal activity of hGIIA in 5448Δ*gacH* and SMUΔ*sccH* expressing catalytically inactive versions of *gacH* and *sccH*, in which the active site T530 and T533 codons were replaced by alanine, respectively (Supplementary Fig. 6). The non-functional *gacH* and *sccH* did not restore hGIIA susceptibility (Fig. 1c,d), indicating that the GroP transferase activity of the *gacH* and *sccH* gene products is required for the observed hGIIA-dependent phenotypes.

GacH cleaves phosphatidylglycerol to release GroP. Experimental evidence suggests that LtaS utilizes the GroP head group of the membrane lipid phosphatidylglycerol as donor for poly(GroP) backbone biosynthesis, liberating diacylglycerol[24,25]. To assess whether GacH also catalyzes the cleavage of phosphatidylglycerol to yield GroP for a transfer reaction, we performed an in vitro experiment employing the eGacH protein and a fluorescently labeled artificial substrate, NBD-phosphatidylglycerol. Incubation of eGacH with NBD-phosphatidylglycerol yielded a fluorescent product (Supplementary Fig. 7a) with the same mobility on silica gel thin-layer chromatography as NBD-diacylglycerol, which was obtained from NBD-phosphatidylglycerol by enzymatic cleavage with phospholipase C from *Bacillus cereus*. Furthermore, the

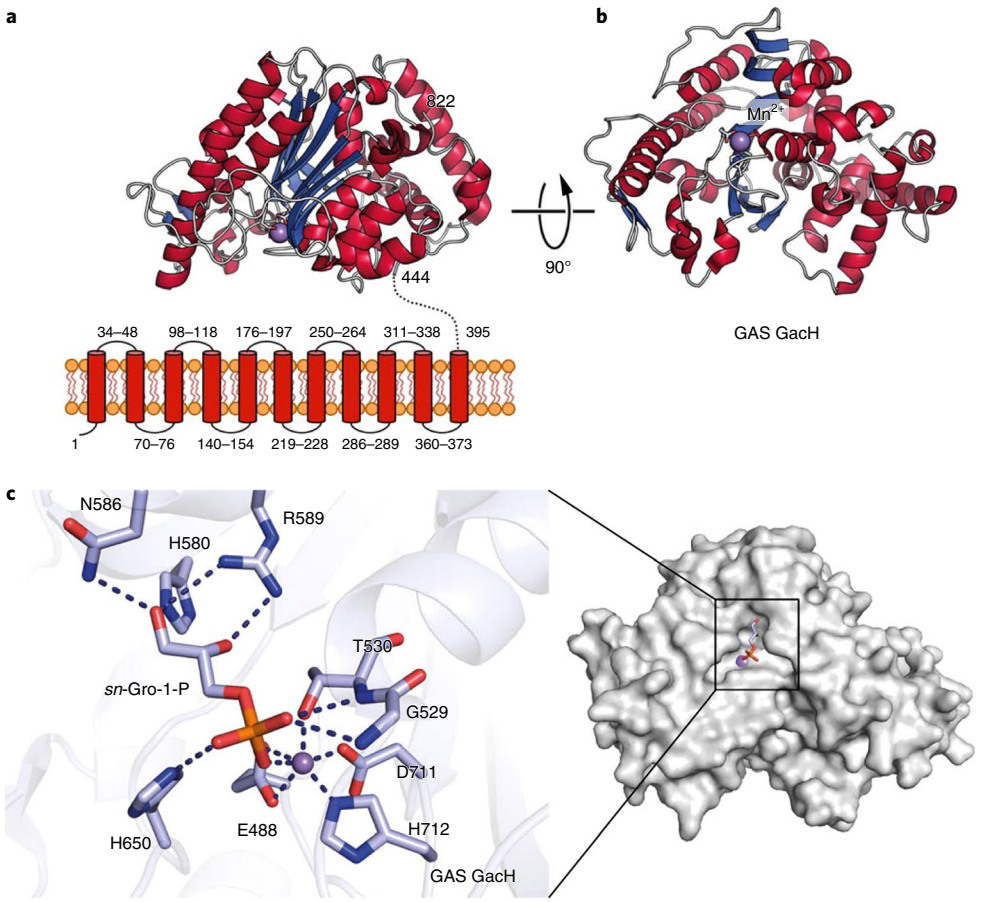


Fig. 3 | Structure of eGacH. a, Predicted topology of GacH showing 11 transmembrane helices and structure of extracellular domain with the enzymatic active site oriented toward the cell membrane. **b**, Structure of apo eGacH viewed at the active site with the Mn^{2+} ion shown as a violet sphere. **c**, A close-up view of the active site GacH crystal structure in complex with *sn*-Gro-1-P.

eGacH product recovered from the silica gel plate and analyzed by liquid chromatography–mass spectrometry (LC–MS) yielded a spectrum consistent with NBD-diacylglycerol and identical to that of the phospholipase C product (Supplementary Fig. 7b–d). The formation of NBD-diacylglycerol by eGacH required the active-site residues, since the eGacH variant, eGacH-T530A, was not active in the assay (Supplementary Fig. 7a). These observations are consistent with the assignment of GacH as a GroP transferase.

GacH homologs decorate glycopolymers with GroP. Phylogenetic analysis of either the full-length or extracellular domains of GacH homologs and LtaS-related proteins revealed that these proteins fall into distinct clades of GroP transferases, suggesting that the proteins may transfer GroP to different substrates (Supplementary Fig. 8). To assess whether *gacH* homologs modify the respective streptococcal glycopolymers with GroP, we enzymatically released GAC and SCC from purified cell walls from GAS and *S. mutans* strains. Subsequently, the enriched polysaccharide preparations were analyzed for glycerol and phosphate. Hydrolysis with HCl released a large amount of glycerol from GAC and SCC isolated from wild-type bacteria (Fig. 4a,b and Supplementary Fig. 9a). Furthermore, we detected high levels of inorganic phosphate after incubation of these acid-treated samples with alkaline phosphatase (Fig. 4a,b and Supplementary Fig. 9a), which was not detected when intact GAC was treated with alkaline phosphatase (Supplementary Fig. 9b,c). This indicates that the phosphoryl moiety is present as a phosphodiester, consistent with its identification as GroP. In contrast to wild-type GAC and SCC, the glycopolymers isolated from 5448Δ*gacH*, 5005Δ*gacH* and SMUΔ*sccH* contained a markedly reduced amount of glycerol and phosphate (Fig. 4a,b and Supplementary Fig. 9a), which was restored only by complementation with wild-type *gacH*, but not *gacH*-T530A, for GAS (Fig. 4a) or plasmid-expressed *sccH* for SMU (Fig. 4b).

In accordance with our functional data, expression of *gacH* did not restore the glycerol and phosphate levels in SCC of SMUΔ*sccH* (Fig. 4b). This suggested that GroP modifications might involve the species-specific side-chains (Glc versus GlcNAc) rather than the identical polyrhamnose backbone. Indeed, the glycerol and phosphate contents in GAC isolated from the GlcNAc-deficient mutant, 5448Δ*gacI*, were significantly reduced (Fig. 4a). Importantly, analysis of GAS strains for total carbohydrate, phospholipid and phosphatidylglycerol contents established that deletion of *gacH* had no effect on these components (Supplementary Fig. 10a–c). Furthermore, the semi-quantitative immuno-dot-blot analysis of GAS strains with anti-GAC antibodies demonstrated that the absolute amount of GAC is not affected by *gacH* deletion (Supplementary Fig. 10d). Analysis of the glycosyl composition of purified cell walls demonstrated that the absence of GacH and SccH did not affect the Rha/GlcNAc and Rha/Glc ratios, respectively (Supplementary Fig. 10e,f).

To provide further evidence that GAC is modified with GroP, GAC samples were subjected to alkaline hydrolysis to release GroP, as described in ref. [26], and the hydrolysate was analyzed by LC–MS for high-molecular weight fragment ions arising from GroP. Compared to wild type, deletion of either *gacH* or *gacI* markedly reduced the levels of GroP in GAC (Supplementary Fig. 11). Complementation of 5448Δ*gacH* with native *gacH*, but not with inactive *gacH*-T530A,

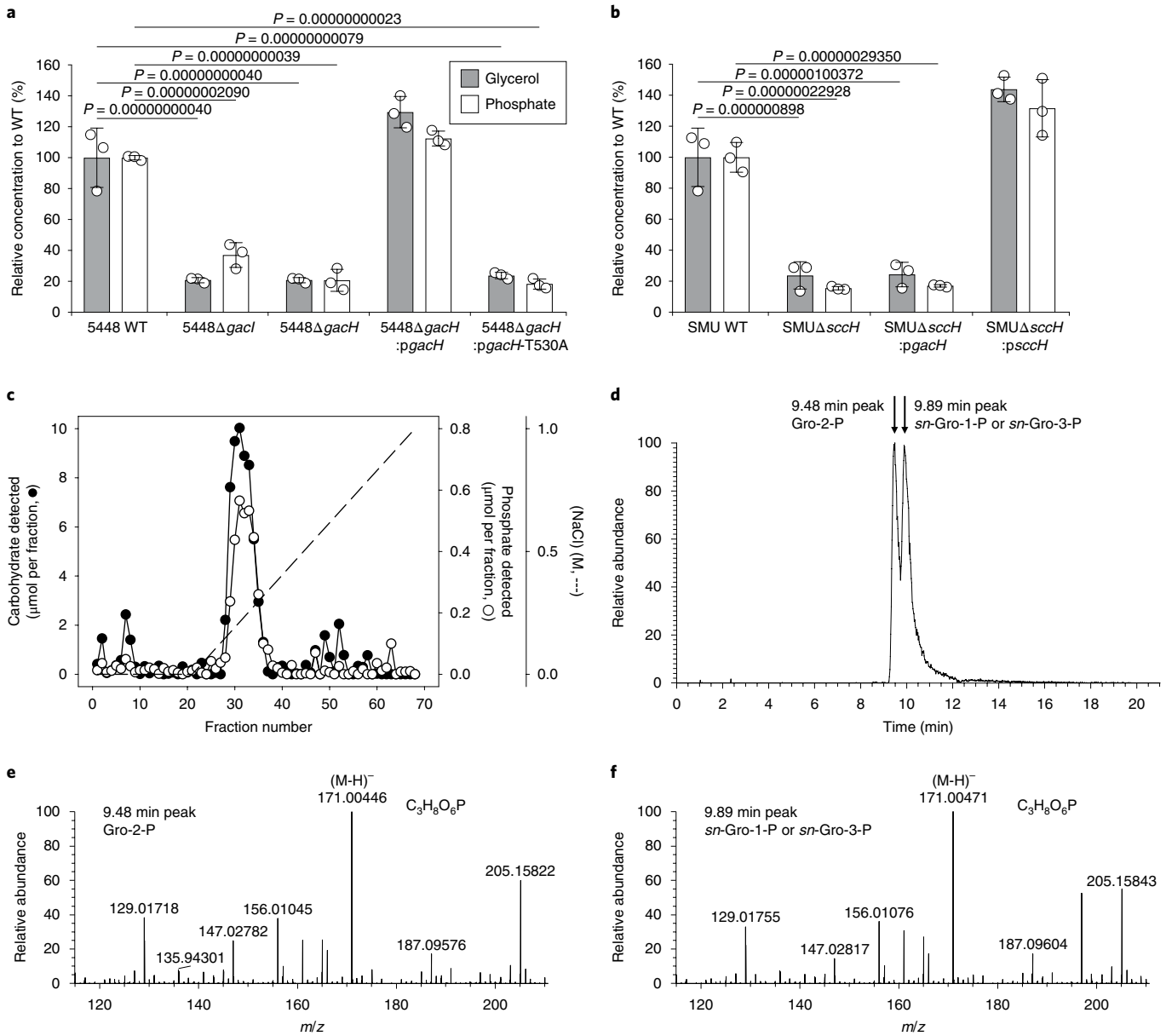


Fig. 4 | GacH and SccH modify their respective glycopolymers with *sn*-Gro-1-P. a,**b**, Analysis of glycerol and phosphate content in GAC and SCC isolated from GAS 5448 wild type, Δ*gacI*, Δ*gacH* and Δ*gacH* complemented with native *gacH* or a catalytically inactivated version of *gacH* (*gacH*-T530A) (**a**), and *S. mutans* wild type, Δ*sccH* and Δ*sccH* complemented with *sccH* or *gacH* (**b**). The concentrations of phosphate and glycerol are presented relative to the wild-type strain. Bars and error bars represent the average and s.d., respectively ($n = 3$ biologically independent samples). *P* values were calculated and adjusted by two-way ANOVA and Bonferroni's multiple comparison test. **c**, Diethylaminoethanol (DEAE)-Sephacel elution profile of GAC isolated from ~90 mg of GAS cell wall. Fractions were analyzed for carbohydrate (●) and phosphate (O). **d**–**f**, Identification of the enantiomeric form of GroP associated with GAC. **d**, The GroP isomers were recovered from GAC following alkaline hydrolysis and separated by liquid chromatography as outlined in Methods. The elution positions corresponding to standard Gro-2-P and *sn*-Gro-1-P/*sn*-Gro-3-P are indicated by the arrows. LC–MS analysis identified two extracted ion chromatogram peaks for the molecular GroP ion mass/charge number (*m/z*) 171.004 $[M\text{-}H]^-$, which eluted at 9.48 (**e**) and 9.89 min (**f**). Based on the accurate mass and retention times, these two peaks were assigned as Gro-2-P and *sn*-Gro-1-P/*sn*-Gro-3-P, respectively, by comparison with authentic chemical standards. Experiments depicted in **c**–**f** were performed independently twice and yielded the same results.

fully restored GroP levels (Supplementary Fig. 11). Thus, the differences in GroP content for the *gacH* and *gacI* deletion mutants are consistent with a role for GacH in modification of the GlcNAc side-chain of GAC with GroP.

To show that GroP is attached directly to GAC, the wild-type GAC was further purified by a combination of size-exclusion and ion-exchange chromatography (Fig. 4c and Supplementary Fig. 12a). The majority of the rhamnose- and phosphate-containing material was bound to the ion-exchange column and eluted as a single coincident peak (Fig. 4c). The GAC purified from 5005Δ*gacH* did not bind to the column (Supplementary Fig. 12b). Interestingly, the 5005Δ*gacH* GAC does appear to contain a small amount of phosphate that may have arisen from the phosphodiester bond linking GAC to peptidoglycan. Taken together, our data directly support the conclusion that GAC is modified with GroP donated by GacH.

GacH decorates GAC with *sn*-Gro-1-P. To assess which GroP enantiomer was incorporated in GAC, GroP liberated from purified GAC by alkaline hydrolysis was analyzed further (Supplementary Fig. 13). As described in detail in ref. [26], if GAC is modified by *sn*-Gro-1-P, alkaline hydrolysis of the phosphodiester bond should result in the formation of a mixture of *sn*-Gro-1-P and Gro-2-P, whereas modification by *sn*-Gro-3-P would yield a mixture of *sn*-Gro-3-P and Gro-2-P[26]. As expected, LC–MS analysis of GAC-derived GroP revealed the presence of two GroP isomers of approximately equal proportions, with LC retention times and major high-molecular weight ions consistent with standard *sn*-Gro-1-P/sn-Gro-3-P and Gro-2-P (Fig. 4d–f and Supplementary Fig. 14). The recovered GroP was characterized further by enzymatic analysis using an *sn*-Gro-3-P assay kit. Under reaction conditions in which *sn*-Gro-3-P standard produced a robust enzymatic signal, incubation with an equal amount of either *sn*-Gro-1-P or the unknown GroP resulted in negligible activity (Supplementary Fig. 15). When *sn*-Gro-3-P was mixed with an equal amount of either *sn*-Gro-1-P or the unknown mixture of GroP isomers, 85.8 and 90.0% of the activity detected with the standard *sn*-Gro-3-P alone was found, confirming that the negative result using the unknown mixture was not due to the presence of an inhibitory compound in GroP preparation. Taken together, our results indicate that GacH decorates GAC with *sn*-Gro-1-P, which is most probably derived from phosphatidylglycerol.

GroP is attached to the C6 hydroxyl group of GlcNAc. To unambiguously establish the presence and location of GroP in GAC, the glycopolymer was isolated from wild-type GAS and analyzed by nuclear magnetic resonance (NMR) (Fig. 5a, Supplementary Table 4 and Supplementary Figs. 16 and 17). The details of NMR analysis are described in Supplementary Notes. Wild type GAC is partially substituted by a GroP residue at O6 of the side-chain β-D-Glc*p*NAc residue; based on integration of the cross-peaks for the anomeric resonances in the ^{1}H,^{13}C-HSQC NMR spectrum, the GAC preparation carries GroP groups to ~25% of the GlcNAc residues. To validate the two-dimensional (2D) NMR results, a triple-resonance ^{1}H,^{13}C,^{31}P NMR experiment based on through-bond $^{1}J_{HC}$ as well as $^{2}J_{CP}$ and $^{3}J_{CP}$ correlations[27] was carried out. The three-dimensional (3D) NMR experiments revealed the ^{1}H NMR chemical shifts of H5′ and the two H6′ protons of the β-D-Glc*p*NAc residue, as well as the two H1 protons and H2 of the Gro residue that all correlated to ^{13}C nuclei (Fig. 5b). The ^{13}C NMR chemical shifts of C5′ and C6′ of the β-D-Glc*p*NAc residue, as well as C1 and C2 of the Gro residue, all correlated to the ^{31}P nucleus (Fig. 5b), and the above protons correlated to the ^{31}P nucleus (Fig. 5b). Taking into consideration the GacH-mediated mechanism of GAC modification by GroP, as well as the biochemical experiments carried out herein, the substituent at O6 of β-D-Glc*p*NAc is an *sn*-Gro-1-P group (Fig. 5c).

Discussion

In Gram-positive bacteria, many peptidoglycan-attached glycopolymers contain negatively charged groups in the repeating units[2]. Previous detailed studies deduced the chemical structure of glycopolymers from GAS and *S. mutans*[3–5]. However, none identified anionic groups in these structures except for one study that reported the presence of glycerol and phosphate in GAC and proposed that this GroP is part of the phosphodiester linkage connecting GAC to peptidoglycan[28]. Similarly, other reports identified substantial concentrations of phosphate in the glycopolymers isolated from a number of streptococcal species[29–31]. Phosphate detection was either disregarded as contamination with LTA[29] or further analyzed using ^{1}H NMR or ^{13}C NMR methods[5,31,32] that do not directly detect phosphoryl moieties in polysaccharides. In our report, we unambiguously confirm that the glycopolymers of GAS and *S. mutans* are in fact polyanionic molecules through decoration of their respective glycan side-chains with GroP (Fig. 5c,d).

We identified and structurally characterized an alternative class of GroP transferase enzymes, represented by GacH, which modifies GAC with GroP in the human pathogen GAS. According to our phylogenetic analysis, GacH homologs are present in many streptococci (Supplementary Fig. 8), suggesting that these bacteria express glycopolymers with GroP-modified side-chains, as we have demonstrated here for *S. mutans*. GacH belongs to the alkaline phosphatase superfamily of which two GroP transferases involved in LTA synthesis, LtaS and LtaP, have been biochemically and structurally characterized[21–23,33]. LtaS and LtaP are membrane proteins that use the membrane lipid phosphatidylglycerol as the GroP donor for the transfer reaction[25]. Our structural analysis of GacH in complex with GroP indicates that the T530 residue participates in the formation of a GroP-enzyme intermediate similar to observations in LtaS, where the GroP molecule is complexed in the active site threonine residue which functions as a nucleophile in phosphatidylglycerol hydrolysis[21–23]. The importance of this residue was also confirmed functionally, since complementation of *gacH* mutant strains with a T530A *gacH* variant could not restore GroP content in GAC and hGIIA sensitivity to wild-type levels. The observations, that the extracellular domain of GacH cleaves phosphatidylglycerol and the GroP in GAC is the *sn*-Gro-1-P enantiomer, strongly suggest that GacH uses phosphatidylglycerol as its donor substrate for the transfer reaction, similar to LtaS (Fig. 5d).

In Gram-positive bacteria, the modification of teichoic acids with D-alanine provides resistance against antibiotics, cationic antimicrobial peptides and small bactericidal enzymes including hGIIA, and affects Mg^{2+} ion scavenging[1,2,34]. It has been assumed that incorporation of positively charged D-alanine into teichoic acids decreases negative bacterial surface charge resulting in reduced initial binding of cationic antimicrobial peptides to the bacterial surface due to ionic repulsion[35,36]. Our study demonstrates that addition of the negatively charged GroP group to glycopolymers protects streptococci from zinc toxicity but also renders bacteria more sensitive to hGIIA activity.

A large body of evidence indicates that phagocytic cells utilize Zn^{2+} intoxication to suppress the intracellular survival of bacteria[37]. Zinc is essential as a key catalytic or structural element for a wide variety of proteins and its concentration needs to be maintained at a specific level, which requires sophisticated systems for uptake and efflux of metal ions. Hence elevated levels of zinc in the cytosol result in cellular toxicity[19], which for GAS is due to inhibition of central carbon metabolism[38]. One mechanism of microbial susceptibility to zinc toxicity is mediated by extracellular competition of Zn^{2+} for Mn^{2+} transport and thereby mediates toxicity by impairing acquisition of Mn^{2+}, the essential nutrient metal[39]. Accordingly, the phenotypes of our GroP and GlcNAc side-chain-deficient mutants could be explained either by 'trapping' of Zn^{2+} in the wild-type cell wall by GroP, or the increased Mn^{2+}-binding capacity of GroP-modified bacterial cell wall. which has been proposed to act as the conduit for the trafficking of mono- and divalent cations to the membrane[34].

Charge-dependent mechanisms probably underlie the increased hGIIA susceptibility of GAS and *S. mutans* expressing GroP-modified glycopolymers. hGIIA is a highly cationic enzyme that catalyzes the hydrolysis of bacterial phosphatidylglycerol[40,41], ultimately leading to bacterial death through lysis. Traversal of this bactericidal enzyme through the Gram-positive cell wall to the plasma membrane is charge-dependent. Indeed, the absence of D-alanine modifications in teichoic acids severely compromises *S. aureus* survival when challenged with hGIIA[41,42]. Similarly, the GacH/SccH-dependent GroP modifications on glycopolymers are required for hGIIA to exert its bactericidal effect against GAS and *S. mutans*, respectively. We have previously demonstrated that loss of the entire GlcNAc GAC side-chain strongly hampers hGIIA trafficking

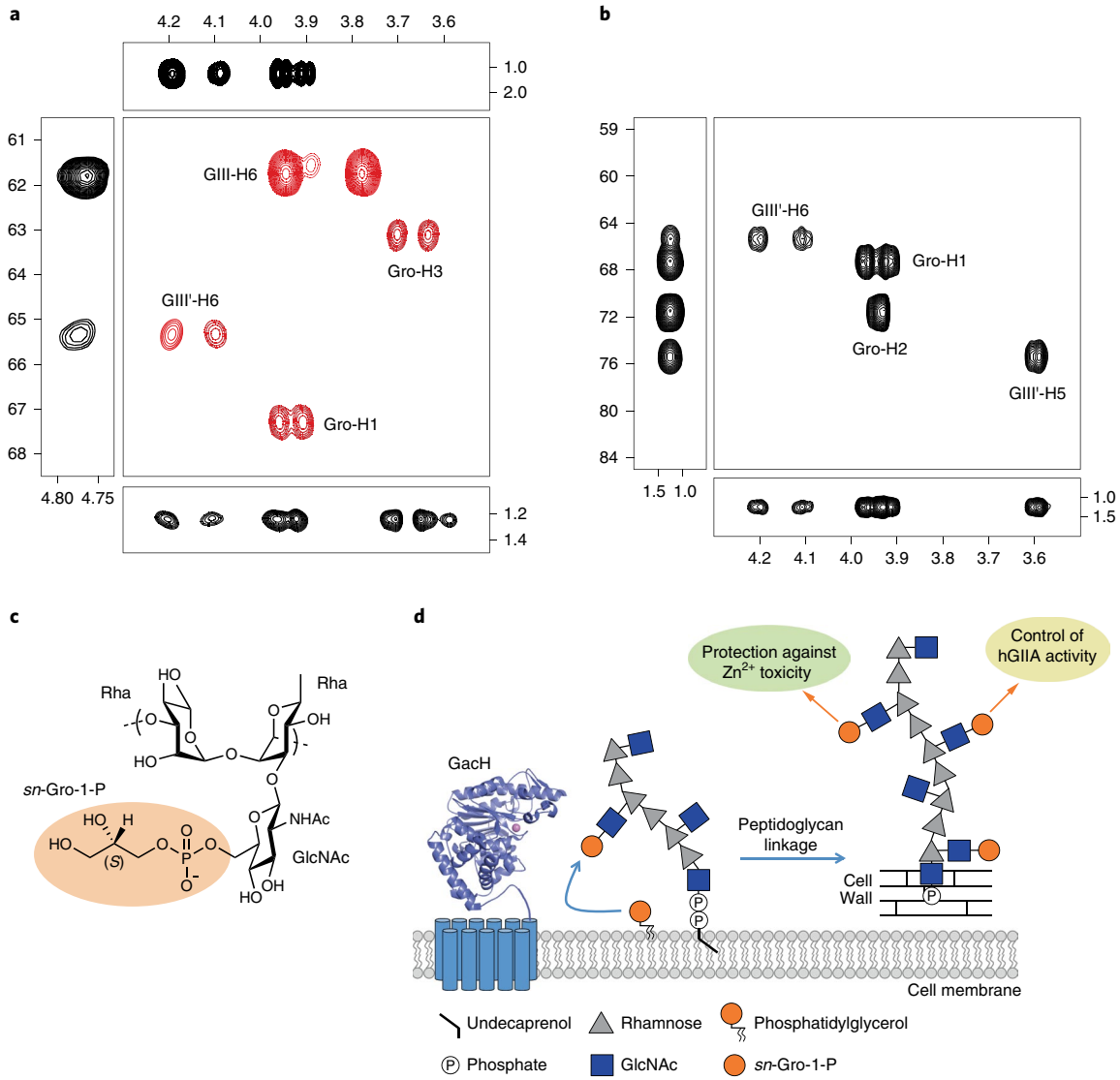


Fig. 5 | NMR analysis confirms the presence of GroP on the C6 GlcNAc hydroxymethyl group of GAC. a,**b**, Selected regions of NMR spectra of GAC. **a**, Multiplicity-edited $^{1}H,^{13}C$-HSQC in which methylene groups have opposite phase and are shown in red (center box); $^{1}H,^{13}C$-HSQC-TOCSY with an isotropic mixing time of 120 ms (left box); $^{1}H,^{13}C$-HMBC with a mixing time of 90 ms (top box); $^{1}H,^{31}P$-hetero-TOCSY with an isotropic mixing time of 80 ms (bottom box). **b**, $^{1}H,^{13}C$-plane (center box); ^{13}C, ^{31}P-plane using a nominal $^{n}J_{CP}$ value of 5 Hz (left box); ^{1}H, ^{31}P-plane (bottom box) of a through-bond 3D $^{1}H,^{13}C,^{13}P$ NMR experiment. Cross-peaks are annotated as GIII corresponding to the GlcNAc residue, GIII′ being the GroP-substituted GlcNAc residue and Gro the glycerol residue. NMR chemical shifts of ^{1}H (horizontal axis), ^{13}C (left axis) and ^{31}P (right axis and left box in **b**) are given in ppm. Experiments depicted in **a**,**b** were performed independently three times and yielded the same results. **c**, Schematic structure of the GAC repeating unit consisting of →3)-α-L-Rha*p*-(1→2)[β-D-Glc*p*NAc6*P*(*S*)Gro-(1→3)]-α-L-Rha*p*-(1→. **d**, The mechanism and roles of GroP cell wall modification in streptococci.

through the GAS cell wall, with a minor contribution of reduced hGIIA binding to the cell surface[17]. Since GroP modifications were also lost in the GlcNAc side-chain-deficient mutant, 5448Δ*gacI*, described in this study, we now assume that the mechanisms of the hGIIA-dependent phenotype are similar in the *gacI* and *gacH* mutants.

Another very important aspect of our study is the identification of an alternative, potentially antigenic, epitope on the surface of streptococci. GAS is associated with numerous mild to life-threatening invasive diseases[43] and is also causative of post-infectious sequelae, including rheumatic heart disease[43]. In particular, the invasive manifestations and post-infectious sequelae are difficult to treat with antibiotics and a GAS vaccine is urgently needed to combat these neglected diseases. The GAC is an attractive candidate for GAS vaccine development, due to its conserved expression in all GAS serotypes and the absence of the constitutive component of GAC, Rha, in humans[12,13]. However, it has been proposed that the GAC GlcNAc side-chain may elicit cross-reactive antibodies relevant to the pathogenesis of rheumatic fever and rheumatic heart disease[44–46]. Moreover, persistence of anti-GAC and anti-GlcNAc antibodies is associated with a poor prognosis in rheumatic heart disease[45]. These clinical associations and the lack of understanding of the pathogenesis of GAS post-infectious rheumatic heart disease have hampered progress in the development of GAC-based vaccines against GAS. However, the GAC GlcNAc decorated with GroP might be an attractive candidate for GAS vaccine development because GroP-modified GlcNAc represents a unique epitope that is absent from human tissues. Thus, our study has implications for the design of a safe and effective vaccine against this important human pathogen for which a vaccine is currently lacking.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41589-019-0251-4.

Received: 8 July 2018; Accepted: 20 February 2019;
Published online: 1 April 2019

References

1. Brown, S., Santa Maria, J. P. Jr. & Walker, S. Wall teichoic acids of Gram-positive bacteria. *Annu. Rev. Microbiol.* **67**, 313–336 (2013).
2. Weidenmaier, C. & Peschel, A. Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nat. Rev. Microbiol.* **6**, 276–287 (2008).
3. Mistou, M. Y., Sutcliffe, I. C. & van Sorge, N. M. Bacterial glycobiology: rhamnose-containing cell wall polysaccharides in Gram-positive bacteria. *FEMS Microbiol. Rev.* **40**, 464–479 (2016).
4. Huang, D. H., Rama Krishna, N. & Pritchard, D. G. Characterization of the group A streptococcal polysaccharide by two-dimensional ^{1}H-nuclear-magnetic-resonance spectroscopy. *Carbohydr. Res.* **155**, 193–199 (1986).
5. St Michael, F. et al. Investigating the candidacy of the serotype specific rhamnan polysaccharide based glycoconjugates to prevent disease caused by the dental pathogen *Streptococcus mutans*. *Glycoconj. J.* **35**, 53–64 (2018).
6. Van der Beek, S. L. et al. A Streptococcus and defines a new class of monomeric dTDP-4-dehydrorhamnose reductases (RmlD). *Mol. Microbiol.* **98**, 946–962 (2015).
7. Tsuda, H., Yamashita, Y., Shibata, Y., Nakano, Y. & Koga, T. Genes involved in bacitracin resistance in *Streptococcus mutans*. *Antimicrob. Agents Chemother.* **46**, 3756–3764 (2002).
8. De, A. et al. Deficiency of RgpG causes major defects in cell division and biofilm formation, and deficiency of LytR-CpsA-Psr family proteins leads to accumulation of cell wall antigens in culture medium by *Streptococcus mutans*. *Appl. Environ. Microbiol.* **83**, e00928 (2017).
9. Nagata, E. et al. Serotype-specific polysaccharide of *Streptococcus mutans* contributes to infectivity in endocarditis. *Oral Microbiol. Immunol.* **21**, 420–423 (2006).
10. Henningham, A. et al. Virulence role of the GlcNAc side chain of the Lancefield cell wall carbohydrate antigen in non-M1-serotype group A *Streptococcus*. *mBio* **9**, e02294 (2018).
11. van Sorge, N. M. et al. The classical Lancefield antigen of group A *Streptococcus* is a virulence determinant with implications for vaccine design. *Cell Host Microbe* **15**, 729–740 (2014).
12. Kabanova, A. et al. Evaluation of a group A *Streptococcus* synthetic oligosaccharide as vaccine candidate. *Vaccine* **29**, 104–114 (2010).
13. Sabharwal, H. et al. group A *Streptococcus* (GAS) carbohydrate as an immunogen for protection against GAS infection. *J. Infect. Dis.* **193**, 129–135 (2006).
14. Shibata, Y., Yamashita, Y., Ozaki, K., Nakano, Y. & Koga, T. Expression and characterization of streptococcal rgp genes required for rhamnan synthesis in *Escherichia coli*. *Infect. Immun.* **70**, 2891–2898 (2002).
15. Rush, J. S. et al. Themolecular mechanism of N-acetylglucosamine side-chain attachment to the Lancefield group A carbohydrate in *Streptococcus pyogenes*.*J. Biol. Chem.* **292**, 19441–19457 (2017).
16. Le Breton, Y. et al. Essential genes in the core genome of the human pathogen *Streptococcus pyogenes*. *Sci. Rep.* **5**, 9838 (2015).
17. van Hensbergen, V. P. et al. Streptococcal Lancefield polysaccharides are critical cell wall determinants for human Group IIA secreted phospholipase A2 to exert its bactericidal effects. *PLoS Pathog.* **14**, e1007348 (2018).
18. Weiss, J. P. Molecular determinants of bacterial sensitivity and resistance to mammalian Group IIA phospholipase A2. *Biochim. Biophys. Acta* **1848**, 3072–3077 (2015).
19. Ong, C. L., GillenC. M., Barnett, T. C., Walker, M. J. & McEwan, A. G. An antimicrobial role for zinc in innate immune defense against group A *Streptococcus*.*J. Infect. Dis.* **209**, 1500–1508 (2014).
20. Graham, M. R. et al. Virulence control in group A *Streptococcus* by a two-component gene regulatorysystem: global expression profiling and in vivo infection modeling. *Proc. Natl Acad. Sci. USA* **99**, 13855–13860 (2002).
21. Lu, D. et al. Structure-based mechanism of lipoteichoic acid synthesis by *Staphylococcus aureus* LtaS. *Proc. Natl Acad. Sci. USA* **106**, 1584–1589 (2009).
22. Schirner, K., Marles-Wright, J., Lewis, R. J. & Errington, J. Distinct and essential morphogenic functions for wall- and lipo-teichoic acids in *Bacillus subtilis*. *EMBO J.* **28**, 830–842 (2009).
23. Campeotto, I. et al. Structural and mechanistic insight into the *Listeria monocytogenes* two-enzyme lipoteichoic acid synthesis system. *J. Biol. Chem.* **289**, 28054–28069 (2014).
24. Fischer, W., Laine, R. A. & Nakano, M. On the relationship between glycerophosphoglycolipids and lipoteichoic acids in Gram-positive bacteria. II Structures of glycerophosphoglycolipids. *Biochim. Biophys. Acta* **528**, 298–308 (1978).
25. Karatsa-Dodgson, M., Wormann, M. E. & Grundling, A. In vitro analysis of the *Staphylococcus aureus* lipoteichoic acid synthase enzyme using fluorescently labeled lipids. *J. Bacteriol.* **192**, 5341–5349 (2010).
26. Kennedy, E. P., Rumley, M. K., Schulman, H. & Van Golde, L. M. Identification of *sn*-glycero-1-phosphate and phosphoethanolamine residues linked to the membrane-derived oligosaccharides of *Escherichia coli*. *J. Biol. Chem.* **251**, 4208–4213 (1976).
27. Marino, J. P. et al. Three-dimensional triple-resonance ^{1}H, ^{13}C, ^{31}P experiment: sequential through-bond correlation of ribose protons and intervening phosphorus along the RNA oligonucleotide backbone. *J. Am. Chem. Soc.* **116**, 6472–6473 (1994).
28. Heymann, H., Manniello, J. M. & Barkulis, S. S. Structure of streptococcal cell walls. V. Phosphate esters in the walls of group A *Streptococcus pyogenes*. *Biochem. Biophys. Res. Commun.* **26**, 486–491 (1967).
29. Emdur, L. I., Saralkar, C., McHugh, J. G. & Chiu, T. H. Glycerolphosphate-containing cell wall polysaccharides from *Streptococcus sanguis*. *J. Bacteriol.* **120**, 724–732 (1974).
30. Prakobphol, A., Linzer, R. & Genco, R. J. Purification and characterization of a rhamnose-containing cell wall antigen of *Streptococcus mutans* B13 (serotype d). *Infect. Immun.* **27**, 150–157 (1980).
31. Pritchard, D. G., Michalek, S. M., McGhee, J. R. & Furner, R. L. Structure of the serotype f polysaccharide antigen of *Streptococcus mutans*. *Carbohydr. Res.* **166**, 123–131 (1987).
32. Pritchard, D. G., Gregory, R. L., Michalek, S. M. & McGhee, J. R. Characterization of the serotype e polysaccharide antigen of *Streptococcus mutans*. *Mol. Immunol.* **23**, 141–145 (1986).
33. Vickery, C. R., Wood, B. M., Morris, H. G., Losick, R. & Walker, S. Reconstitution of *Staphylococcus aureus* lipoteichoic acid synthase activity identifies Congo red as a selective inhibitor. *J. Am. Chem. Soc.* **140**, 876–879 (2018).
34. Neuhaus, F. C. & Baddiley, J. A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in Gram-positive bacteria. *Microbiol. Mol. Biol. Rev.* **67**, 686–723 (2003).
35. Peschel, A. et al. Inactivation of the *dlt* operon in *Staphylococcus aureus* confers sensitivity to defensins, protegrins, and other antimicrobial peptides. *J. Biol. Chem.* **274**, 8405–8410 (1999).
36. Falagas, M. E., Rafailidis, P. I. & Matthaiou, D. K. Resistance to polymyxins: mechanisms, frequency and treatment options. *Drug Resist. Updat.* **13**, 132–138 (2010).
37. Djoko, K. Y., Ong, C. L., Walker, M. J. & McEwan, A. G. The role of copper and zinc toxicity in innate immune defense against bacterial pathogens. *J. Biol. Chem.* **290**, 18954–18961 (2015).
38. Ong, C. L., Walker, M. J. & McEwan, A. G. Zinc disrupts central carbon metabolism and capsule biosynthesis in *Streptococcus pyogenes*. *Sci. Rep.* **5**, 10799 (2015).
39. McDevitt, C. A. et al. A molecular mechanism for bacterial susceptibility to zinc. *PLoS Pathog.* **7**, e1002357 (2011).
40. Buckland, A. G. & Wilton, D. C. Inhibition of secreted phospholipases A2 by annexin V. Competition for anionic phospholipid interfaces allows an assessment of the relative interfacial affinities of secreted phospholipases A2. *Biochim. Biophys. Acta* **1391**, 367–376 (1998).
41. Koprivnjak, T., Peschel, A., Gelb, M. H., Liang, N. S. & Weiss, J. P. Role of charge properties of bacterial envelope in bactericidal action of human group IIA phospholipase A2 against *Staphylococcus aureus*. *J. Biol. Chem.* **277**, 47636–47644 (2002).
42. Hunt, C. L., Nauseef, W. M. & Weiss, J. P. Effect of d-alanylation of (lipo) teichoic acids of *Staphylococcus aureus* on host secretory phospholipase A2 action before and after phagocytosis by human neutrophils. *J. Immunol.* **176**, 4987–4994 (2006).
43. Carapetis, J. R., Steer, A. C., Mulholland, E. K. & Weber, M. The global burden of group A streptococcal diseases. *Lancet Infect. Dis.* **5**, 685–694 (2005).
44. Goldstein, I., Rebeyrotte, P., Parlebas, J. & Halpern, B. Isolation from heart valves of glycopeptides which share immunological properties with *Streptococcus haemolyticus* group A polysaccharides. *Nature* **219**, 866–868 (1968).
45. Ayoub, E. M. & Dudding, B. A. Streptococcal group A carbohydrate antibody in rheumatic and nonrheumatic bacterial endocarditis. *J. Lab. Clin. Med.* **76**, 322–332 (1970).
46. Kirvan, C. A., Swedo, S. E., Heuser, J. S. & Cunningham, M. W. Mimicry and autoantibody-mediated neuronal cell signaling in Sydenham chorea. *Nat. Med.* **9**, 914–920 (2003).
47. Le Breton, Y. et al. Genome-wide discovery of novel M1T1 group A streptococcal determinants important for fitness and virulence during soft-tissue infection. *PLoS Pathog.* **13**, e1006584 (2017).
48. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
49. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

Acknowledgements

This work was supported by the Center of Biomedical Research Excellence (COBRE) Pilot Grant (to N.K., K.V.K. and J.S.R.) supported by NIH grant No. P30GM110787 from the National Institute of General Medical Sciences (NIGMS); NIH grant No. R56AI135021 from the National Institute of Allergy and Infectious Diseases (NIAID) (to N.K.); VIDI grant No. 91713303 from the Netherlands Organization for Scientific Research (NWO) (to N.M.vS. and V.P.vH.); the Swedish Research Council (Nos. 2013–4859 and 2017–03703) and The Knut and Alice Wallenberg Foundation (to G.W.); NIH grant No. P30GM127211 from the NIGMS and NIH grant No. 1S10OD021753 (to A.J.M.); the National Health and Medical Research Council of Australia (to M.J.W.); grants from CNRS, ANR (MNaims No. ANR-17-CE17-0012-01) and FRM (No. SPF20150934219) (to G.L.); NIH grant No. AI047928 from NIAID (to K.S.M. and Y.L.B.); and NIH grant No. AI094773 (to N.M.E.S. and A.T.B.). Carbohydrate composition analysis at the Complex Carbohydrate Research Center was supported by the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, US Department of Energy grant (No. DE-FG02-93ER20097) to P.A. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31–109-Eng-38 and NIH grants Nos. S10_RR25528 and S10_RR028976. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the NIH, NIGMS including No. P41GM103393. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH.

Author contributions

A.R., P.D., Y.L.B., K.S.M., A.G.M., A.J.M., G.L., M.J.W., J.S.R., K.V.K., G.W., N.M.vS. and N.K. designed the experiments. R.J.E., V.P.vH., A.R., A.T., J.S.R., K.V.K., G.W. and N.K. performed functional and biochemical experiments. K.V.K. carried out X-ray crystallography and structure analysis. A.R. and G.W. performed NMR studies. P.D. and A.J.M. performed MS analysis. V.P.vH., N.K. and K.V.K. constructed plasmids and isolated mutants. R.J.E., V.P.vH., A.R., P.D., Y.L.B., N.M.E.S., A.T.B., K.S.M., A.G.M., A.J.M., M.J.W., J.S.R., K.V.K., G.W., N.M.vS. and N.K. analyzed the data. N.M.vS. and N.K. wrote the manuscript with contributions from all authors. All authors reviewed the results and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41589-019-0251-4.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to N.M.v. or N.K.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Bacterial strains, growth conditions and media. All plasmids, strains and primers used in this study are listed in Supplementary Tables 5 and 6. GAS and *S. mutans* strains were grown in Todd–Hewitt broth supplemented with 1% yeast extract (THY) without aeration at 37 °C. *S. mutans* plates were grown with 5% CO_2. For hGIIA-mediated killing experiments, *S. mutans* strains were grown in Todd–Hewitt broth without yeast extract and with 5% CO_2. *E. coli* strains were grown in Lysogeny broth medium or on Lysogeny broth 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 (Erm) at 500 μg ml^{-1} for *E. coli*, 5 μg ml^{-1} for GAS and 10 μg ml^{-1} for *S. mutans*; chloramphenicol (CAT) at 10 μg ml^{-1} for *E. coli* and 2 μg ml^{-1} for GAS and *S. mutans*; spectinomycin at 200 μg ml^{-1} for *E. coli*, 100 μg ml^{-1} for GAS and 500 μg ml^{-1} for *S. mutans*; and kanamycin at 300 μg ml^{-1} for GAS.

To identify genes providing resistance against Zn^{2+} toxicity, Roswell Park Memorial Institute (RPMI) 1640 (without glucose) (Gibco) was supplemented with guanine, adenine and uracil at a concentration of 25 μg ml^{-1} each, D-glucose at a concentration of 0.5% w/v and HEPES at 50 mM. Vitamins were provided by BME Vitamins 100× (Sigma, No. B6891).

Genetic manipulations. Plasmids were transformed into GAS and *S. mutans* by electroporation or natural transformation as described previously[6]. All constructs and mutants were confirmed by PCR and sequencing analysis (Eurofins MWG Operon and Macrogen).

Genetic manipulation of GAS: To construct 5005Δ*gacH* and 2221Δ*gacH*, 5005 chromosomal DNA was used as a template for amplification of two DNA fragments using primer pairs 5005-f/gacHdel-r and gacHdel-f/5005-r, which were fused and amplified using a PCR overlap method[50] with primer pair 5005-f/5005-r to create the deletion of *gacH*. The PCR product was digested with BamHI and XhoI and ligated into BamHI/SalI-digested plasmid pBBL740, transformed into 5005 and 2221, and CAT-resistant colonies were selected on THY agar plates. Several potential double-crossover mutants were selected as previously described[51].

To construct the plasmid for in cis complementation of 5005Δ*gacH*, 5005 chromosomal DNA was used as a template for amplification of *gacH* using the primer pair 5005-f/5005-r, which was cloned in pBBL740 through restriction-ligation using BamHI and XhoI. The plasmid was transformed into the 5005Δ*gacH* strain, and CAT-resistant colonies were selected on THY agar plates. Double-crossover mutants were selected as described above.

To construct 5448Δ*gacH*, GAS 5448 chromosomal DNA was used to amplify up- and downstream regions flanking *gacH* using primer pairs 5448-f/5448CAT-r and 5448CAT-f/5448-r. Primers 5448CAT-f and 5448CAT-r contain 25 base pair (bp) extensions complementary to the CAT resistance cassette. Up- and downstream were fused to the CAT cassette using 5448-f/5448-r, and cloned into pHY304 through digestion-ligation using XhoI and HindIII, yielding plasmid pHY304Δ*gacH*. After plasmid transformation into 5448, the double-crossover mutant 5448Δ*gacH* was selected as previously described[15].

To complement 5448Δ*gacH*, *gacH* was amplified from 5448 chromosomal DNA using primer pair gacH-EcoRI-f/gacH-BglII-r, digested using EcoRI/BglII and ligated into EcoRI/BglII-digested pDCerm, yielding p*gacH*_erm. To make a catalytically inactive variant of *gacH*, the mutation T550A was introduced into p*gacH*_erm using Gibson Assembly site-directed mutagenesis with the primers gacH-T530A-F, gacH-T530A-R, repB-isoF and repB-isoR. The plasmids were transformed into 5448Δ*gacH* and selected for Erm resistance on THY agar plates. Transformation was confirmed by PCR, yielding strains *5448*Δ*gacH*:p*gacH* and *5448*Δ*gacH*:p*gacH*-T530A, respectively.

To construct SMUΔ*sccH*, *S. mutans Xc* chromosomal DNA was used to amplify up- and downstream region flanking using primer pairs sccH-f/sccH-erm-r and sccH-erm-f /sccH-r. Primers sccH-erm-f and sccH-erm-r contained 25 bp extensions complementary to the Erm resistance cassette. Up- and downstream PCR fragments were mixed with the Erm cassette and amplified as a single PCR fragment using primer pair sccH-f/sccH-r. The *sccH* knockout construct was transformed into *S. mutans* as described previously[6]. Erm-resistant single colonies were picked and checked for deletion of *sccH* and integration of Erm cassette by PCR, resulting in SMUΔ*sccH*. For complementation, *sccH* and *gacH* were amplified from *S. mutans Xc* and GAS 5448 chromosomal DNA, respectively, using primer pairs sccH-EcoRI-f/sccH-BglII-r and gacH-EcoRI-f/gacH-BglII-r. The PCR products were digested with EcoRI/BglII and ligated into EcoRI/BglII-digested pDC123 vector, yielding p*sccH* and p*gacH*_cm, respectively. To make a catalytically inactive variant of *sccH*, the mutation T553A was introduced into p*sccH* using Gibson Assembly site-directed mutagenesis with the primers sccH-T553A-F, sccH-T553-R, repB-isoF and repB-isoR. The plasmids were transformed into SMUΔ*sccH* as previously described[6]. CAT-resistant single colonies were picked and checked for the presence of p*sccH* or p*gacH*_cm by PCR, yielding strains SMUΔ*sccH*:p*sccH*, SMUΔ*sccH*:p*sccH*-T533A and SMUΔ*sccH*:p*gacH*, respectively.

To create a vector for expression of eGacH in *E. coli*, the gene was amplified from 5005 chromosomal DNA using the primers gacH-NcoI-f and gacH-XhoI-r. The PCR product was digested with NcoI and XhoI, and ligated into NcoI/XhoI-digested pCDF-NT vector. The resultant plasmid, pCDF-GacH, contained *gacH* fused at the N terminus with a His-tag followed by a TEV protease recognition site. To produce a catalytically inactive variant of eGacH, the mutation T530A was introduced into pCDF-GacH using Gibson Assembly site-directed mutagenesis and the primers gacH-T530A-F, gacH-T530A-R, Str-isoF and Str-isoR.

Identification of hGIIA-resistant GAS transposon mutants. The GAS M1T1 5448 *Krmit* transposon mutant library[16] was grown to mid-log phase (OD_{600} = 0.4). Colony-forming units (1×10^5) were incubated with 27.5 μg ml^{-1} recombinant hGIIA[52] in triplicate for 1 h at 37 °C and plated on THY agar plates supplemented with kanamycin. The position of the transposon insertion of resistant colonies was determined as described previously[53]. hGIIA susceptibility experiments were performed as described previously[17].

Determination of selective metal concentrations. To determine the target concentration of Zn^{2+}, colonies of 5448 wild type and 5448Δ*czcD*[19] were scraped from THY agar plates, resuspended and washed in phosphate buffered saline (PBS) to OD_{600} = 1 and used to inoculate freshly prepared modified RPMI (mRPMI) containing varying concentrations of Zn^{2+} to OD_{600} = 0.05 in a 96-well plate. Growth at 37 °C was monitored at OD_{595} every 15 min using the BMG Fluostar plate reader.

Tn-seq library screen for Zn^{2+} sensitivity. The 5448 *Krmit* Tn-seq library at T_0 generation[16] was thawed, inoculated into 150 ml prewarmed THY broth containing kanamycin and grown at 37 °C for 6 h. The culture (T_1) was centrifuged at 4,000*g* for 15 min at 4 °C and the pellet resuspended in 32.5 ml saline. Freshly prepared mRPMI or mRPMI containing 10 μM or 20 μM Zn^{2+} was inoculated with 500 μl culture into 39.5 ml media, creating a 1:20-fold inoculation. These T_2 cultures were then grown at 37 °C for exactly 6 h, at which point 2 ml of these cultures was inoculated again into 38 ml of freshly prepared mRPMI alone or mRPMI containing 10 μM or 20 μM Zn^{2+}. The remaining 38 ml of T_2 culture was harvested by centrifugation at 4,000*g* for 10 min at 4 °C and pellets stored at −20 °C for subsequent DNA extraction. Cultures were grown for additional 6 h, at which point T_3 cultures were harvested by centrifugation at 4,000*g* for 10 min at 4 °C and pellets stored at −20 °C.

Tn-seq *Krmit* transposon insertion tags were prepared from the cell pellets as previously described[16,47]. After quality control with the Bioanalyzer instrument (Agilent), the libraries of *Krmit* insertion tags were sequenced (50 nt single-end reads) on an Illumina HiSeq 1500 in the Institute for Bioscience and Biotechnology Research Sequencing Core at the University of Maryland, College Park. Tn-seq read datasets were analyzed (quality, filtering, trimming, alignment, visualization) as previously described[16,47] using the M1T1 5448 genome as reference for read alignments. The ratios of mutant abundance comparing the output to input mutant pools were calculated as a fold change for each GAS gene using the DEseq2 and EdgeR pipelines[47–49].

Drop-test assays. Strains 5448 wild type, 5448Δ*gacI*, 5448Δ*gacI*:*gacI*, 5448Δ*gacH*, 5448Δ*gacH*:p*gacH*, *S. mutans* wild type, SMUΔ*sccH*, SMUΔ*sccH*:p*sccH* and SMUΔ*sccH*:p*gacH* were grown in THY to mid-exponential growth phase, adjusted to OD_{600} = 0.6 and serially diluted. Five microliters were spotted onto THY agar plates containing varying concentrations of Zn^{2+} ($ZnSO_4 \cdot 7H_2O$). Plates were incubated at 37 °C overnight and photographed.

Protein expression and purification. To purify eGacH and eGacH-T530A, *E. coli* Rosetta(DE3) carrying the respective plasmid was grown in Lysogeny broth at 37 °C to OD_{600} = 0.4–0.6 and induced with 0.25 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 18 °C for approximately 16 h. Bacteria were lysed in 20 mM Tris-HCl pH 7.5, 300 mM NaCl by a microfluidizer cell disrupter. The soluble fractions were purified by nickel-nitrilotriacetic acid chromatography. The eluate was dialyzed into 20 mM Tris-HCl pH 7.5, 300 mM NaCl in the presence of TEV protease (1 mg 20 mg^{-1} protein) and reapplied to a nickel-nitrilotriacetic acid column equilibrated in 20 mM Tris-HCl pH 7.5, 300 mM NaCl to remove the cleaved His-tag and any uncleaved protein from the sample. The protein was further purified by size-exclusion chromatography on a Superdex 200 column in 20 mM HEPES pH 7.5, 100 mM NaCl.

To express seleno-methionine (Se-Met)-labeled eGacH, *E. coli* Rosetta(DE3) carrying eGacH was grown in Lysogeny broth at 37 °C to OD_{600} = 0.5. The bacteria were centrifuged and resuspended in M9 minimal media supplemented with Se-Met. Protein expression was induced with 0.25 mM IPTG, and the cultures were grown at 16 °C for approximately 16 h. Se-Met-labeled eGacH was purified as described above.

Crystallization, data collection and structure solution. The conditions for eGacH crystallization were initially screened using the JCSG Suites I–IV screens (Qiagen) at a protein concentration of 9 mg ml^{-1} by the hanging-drop vapor diffusion method. Crystals of Se-Met-substituted eGacH were grown in 0.1 M HEPES pH 7.5, 10% PEG8000 and 8% ethylene glycol. Crystals were transferred to crystallization solution supplemented with 20% ethylene glycol and flash-frozen in liquid nitrogen. The data were collected at APS 22-ID at a wavelength of 0.9793 Å. Crystals of GroP●eGacH complex were obtained using crystallization solution containing 0.2 M calcium acetate, 0.1 M MES pH 6.0 and 20% PEG8000.

sn-Gro-1-P (Sigma Aldrich) was mixed with eGacH at 10 mM before crystallization. Initial crystals of the GroP•eGacH complex belonged to the same crystal form as apo GacH, but crystals of different morphology grew epitaxially after several days. These crystals displayed better diffraction and were used for structure determination of the GroP•eGacH complex. Crystals were cryoprotected in crystallization solution supplemented with 10 mM *sn*-glycerol-1-phosphate and 20% ethylene glycol and vitrified in liquid nitrogen. The data were collected at SSRL BL9–2 at a wavelength of 0.97946 Å.

All data were processed and scaled using *XDS* and *XSCALE*[54]. The structure of eGacH was solved by the Se single-wavelength anomalous diffraction method. Se atom positions were determined using the HySS module in *PHENIX*[55]. The structure was solved using AutoSol wizard in *PHENIX*[55]. The model was completed using *Coot*[56] and refined using *phenix.refine* in *PHENIX*[55]. Ramachandran distribution analysis of the eGacH final structure with MolProbity[57] indicates that 96.6 and 3.4% residues are in favored and allowed regions, respectively, with no outliers.

The structure of the GroP•eGacH complex was solved by molecular replacement using *Phaser* in *PHENIX*[55] and the dimer of apo eGacH as a search model. The model was adjusted using *Coot* and refined using *phenix.refine*. Difference electron density corresponding to GroP molecules was readily identified after refinement. GroP molecules were modeled using *Coot*. The geometric restraints for GroP were generated using Grade Web Server (http://grade.globalphasing.org) (Global Phasing). The last several rounds of refinement were performed using 19 translation/libration/screw groups, which were identified by *PHENIX*[55]. Ramachandran distribution analysis of the GroP•eGacH final structure indicated that 97.2 and 2.8% residues are in favored and allowed regions, respectively, with no outliers. The structures were validated using *Coot*, MolProbity and wwPDB Validation Service (https://validate.wwpdb.org). Statistics for data collection, refinement and model quality are listed in Supplementary Table 7.

In vitro assay of eGacH enzymatic activity. Purification of 16:0–6:0 NBD-phosphatidylglycerol lipid (Avanti) was performed by preparative thin-layer chromatography as described previously[25], dissolved in CH_3OH and stored at −20 °C until use. The lipid was dried and dispersed in octyl-glucoside by sonication before addition of the remaining components. Reaction mixtures contained 0.05 M sodium succinate pH 6.3, 10 mM $MnCl_2$, 0.05 M NaCl, 0.25% octyl-glucoside, 20 µg NBD-phosphatidylglycerol, ultrasonically dispersed in 0.5% octyl-glucoside (Branson 2200 bath sonicator) and either no enzyme, 20 µg eGacH or 20 µg eGacH-T530A in a total volume of 0.02 ml. Following incubation at 37 °C for 3 h, the reaction was stopped by the addition of 0.08 ml $CHCl_3/CH_3OH$ (2:1) and analyzed for fluorescence on a BioRad ChemiDoc MP Imaging System using the fluorescein preset mode, as described previously[25]. The migration position of the NBD-diacylglycerol product was determined from the product of a separate reaction containing purified phospholipase C from *B. cereus* (Sigma Aldrich).

Isolation of cell wall. Cell wall was isolated from exponential phase cultures by the sodium dodecyl sulfate-boiling procedure and lyophilized as previously described[15].

GAC purification. GAC was released from the cell wall by sequential digestion with mutanolysin (Sigma Aldrich) and recombinant PlyC amidase[15], and partially purified by a combination of size-exclusion chromatography and ion-exchange chromatography. Mutanolysin digests contained 5 mg ml^{-1} of cell wall suspension in 0.1 M sodium acetate, pH 5.5, 2 mM $CaCl_2$ and 5 U ml^{-1} mutanolysin. Following overnight incubation at 37 °C, GAC was separated from the cell wall by centrifugation at 13,000*g*, 10 min and precipitated from 80% acetone (−20 °C). The precipitate was sedimented (5,000*g*, 20 min), dried briefly under nitrogen gas and redissolved in 0.1 M Tris-HCl, pH 7.4 and digested with PlyC (50 µg ml^{-1}) overnight at 37 °C. Following PlyC digestion, GAC was recovered by acetone precipitation as described above, redissolved in a small volume of 0.2 N acetic acid and chromatographed on a 25 ml column of BioGel P10 equilibrated in 0.2 N acetic acid. Fractions (1.5 ml) were collected and monitored for carbohydrate by the anthrone assay. Fractions containing GAC were combined, concentrated and desalted by spin column centrifugation (3,000 MW cutoff filter). GAC was loaded onto an 18 ml column of DEAE-Sephacel. The column was eluted with a 100 ml gradient of NaCl (0–1 M). Fractions were analyzed for carbohydrate by the anthrone assay and phosphate by the malachite green assay following digestion with 70% perchloric acid. Fractions containing peaks of carbohydrate were combined, concentrated by spin column (3,000 MW cutoff) and lyophilized.

Anthrone assay. Total carbohydrate content was determined by a minor modification of the anthrone procedure. Reactions contained 0.08 ml of aqueous sample and water and 0.32 ml anthrone reagent (0.2% anthrone in concentrated H_2SO_4). The samples were heated to 100 °C for 10 min, cooled in water (room temperature) and the absorbance at 580 nm was recorded. GAC concentration was estimated using an L-rhamnose standard curve.

Phosphate assay. Approximately 1.5 mg of GAS cell wall material was dissolved in 400 µl H_2O and 8 µg ml^{-1} PlyC and incubated at 37 °C, rotating for approximately 16 h. Additional PlyC was added and incubated for a further 4–6 h. To liberate SCC from *S. mutans* cell walls, 1.5 mg of *S. mutans* cell wall material was incubated 24 h with 1.5 U ml^{-1} mutanolysin in 400 µl of 0.1 M sodium acetate, pH 5.5 and 2 mM $CaCl_2$. The samples were incubated at 100 °C for 20 min and centrifuged for 5 min at maximum speed in a table top centrifuge. The supernatant was transferred to a new micro-centrifuge tube and incubated with 2 N HCl at 100 °C for 2 h. The samples were neutralized with NaOH in the presence of 62.5 mM HEPES pH 7.5. To 100 µl of acid hydrolyzed sample, 2 µl of 1 U µl^{-1} alkaline phosphatase (Thermo Fisher) and 10 µl 10× alkaline phosphatase buffer was added and incubated at 37 °C, rotating, overnight. Released phosphate was measured using the Pi ColorLock Gold kit (Innova Biosciences), according to the manufacturer's protocol.

Total phosphate content was determined by the malachite green method following digestion with perchloric acid. Samples containing 10–80 µl were heated to 110 °C with 40 µl 70% perchloric acid (Fisher Scientific) in 13 × 100 borosilicate disposable culture tubes for 1 h. The reactions were diluted to 160 µl with water and 100 µl was transferred to a flat-bottom 96-well culture plate. Malachite green reagent (0.2 ml) was added and the absorbance at 620 nm was read after 10 min at room temperature. Malachite green reagent contained 1 vol. 4.2% ammonium molybdate tetrahydrate (by weight) in 4 M HCl, 3 vol. 0.045% malachite green (by weight) in water and 0.01% Tween 20.

Glycerol assay. Samples for glycerol measurement were prepared as described for the phosphate assay but were not digested with alkaline phosphatase. Instead, glycerol concentration was measured using the Glycerol Colorimetric assay kit (Cayman Chemical) according to the manufacturer's protocol.

Glycosyl composition analysis. Glycosyl composition analysis of GAS and *S. mutans* cell wall samples was performed at the Complex Carbohydrate Research Center by combined gas chromatography/mass spectrometry of the per-O-trimethylsilyl derivatives of the monosaccharide methyl glycosides produced from the sample by acidic methanolysis as described previously[15].

Dot-blot analysis of GAC. The semi-quantitative immuno-dot-blot analysis of GAC expressed by GAS strains was conducted essentially as described[15] with the following minor modification: the PlyC-digested cell wall fractions were serially diluted before spotting to a nitrocellulose membrane.

Total phospholipid content analysis. GAS cells (50 ml) grown in THY broth to $OD_{600} = 0.5$ were centrifuged, washed with PBS, resuspended in 5 ml PBS and incubated with PlyC (100 µg ml^{-1}) at 37 °C for 1 h. A total lipid extract was prepared by a modification of Bligh–Dyer extraction, freed of non-lipid contaminants by Folch partitioning, as described previously[15], and redissolved in 1 ml $CHCl_3/CH_3OH$ (2:1). An aliquot (20 µl) was analyzed for total phospholipid using the malachite green phosphate assay following perchloric acid digestion at 120 °C for 1 h, and the remainder was analyzed for phosphatidylglycerol.

Phosphatidylglycerol and NBD-diacylglycerol analysis. The analysis of phosphatidylglycerol and NBD-diacylglycerol was performed using an Ultimate 3000 ultra HPLC system coupled to a Thermo Q-Exactive Orbitrap mass spectrometer equipped with a heated electrospray ion source (Thermo Scientific). Lipid extracts were separated on a Waters ACQUITY BEH C8 column (2.1 × 100 mm, 1.7 µm) with the temperature maintained at 40 °C. The flow rate was 250 µl min^{-1}, and the mobile phases consisted of 60:40 water/acetonitrile (A) and 90:10 isopropanol/acetonitrile (B), both containing 10 mM ammonium formate and 0.1% formic acid. The samples were eluted with a linear gradient from 32% B to 97% B over 25 min, maintained at 97% B for 4 min and re-equilibration with 32% B for 6 min. The sample injection volume was 5 µl. The mass spectrometer was operated in positive and negative ionization modes. The full-scan and fragment spectra were collected at a resolution of 70,000 and 17,500, respectively. Data analysis and lipid identification were performed using Xcalibur 4.0 and Lipidsearch 4.1 (Thermo Fisher).

Total carbohydrate content analysis. GAS cells (10 ml) grown in THY broth $OD_{600} = 0.5$ were centrifuged, washed twice with PBS and resuspended in 0.2 ml distilled water. The cell suspension was assayed for carbohydrate content by the anthrone assay as described above.

Analysis of GAC for GroP. GAC (prepared from ~1.5 mg of cell wall) was hydrolyzed in 0.1 ml 0.5 M NaOH (100 °C, 1 h) to release GroP, as described in ref. [26]. Following alkaline treatment, the reaction was neutralized with acetic acid, supplemented with 5 nmol of citronellyl phosphate (as internal standard) and centrifuged on an Amicon Ultra Centrifugal Filter (0.5 ml, 3,000 nominal molecular weight limit). The filtrate was lyophilized and analyzed with a Q-Exactive mass spectrometer and an Ultimate 3000 ultra HPLC system (Thermo Fisher Scientific) using a silica-based SeQuant ZIC-pHILIC column (2.1 × 150 mm, 5 µm, Merck) with elution buffers consisting of (1) 20 mM $(NH_4)_2CO_3$ with 0.1% NH_4OH in H_2O and (2) acetonitrile. The column temperature was maintained at 40 °C, and the flow rate was set to 150 µl min^{-1}. MS detection was performed by electrospray ionization in negative ionization mode with source voltage maintained at 3.0 kV. The capillary temperature, sheath gas flow and auxiliary gas flow were set

at 275 °C, 40 arb and 15 arb units, respectively. Full-scan MS spectra (mass range m/z 75–1,000) were acquired with resolution $R = 70{,}000$ and AGC target 1×10^6. Extracted ion chromatograms for GroP and citronellyl phosphate were obtained from the LC–MS chromatograms and used to estimate relative GroP content.

Identification of the stereochemistry of the GroP moiety of GAC. GroP was liberated from GAC by alkaline hydrolysis as described in ref. [26] and re-fractionated on BioGel P10. The bulk of the GAC elutes in the void volume and GroP elutes in the inclusion volume as identified by LC–MS. Column fractions containing GroP were combined, concentrated by rotary evaporation (30 °C, under reduced pressure) and desalted on BioGel P2. The stereochemistry of the GroP was determined by an enzymatic method using the Amplite Fluorimetric *sn*-Gro-3-P Assay Kit (AAT Bioquest) according to the manufacturer's instructions.

NMR spectroscopy. The NMR spectra were recorded on a Bruker AVANCE III 700 MHz equipped with a 5 mm TCI Z-Gradient Cryoprobe (^{1}H/^{13}C/^{15}N) and dual receivers, and a Bruker AVANCE II 600 MHz spectrometer equipped with a 5 mm TXI inverse Z-Gradient ^{1}H/D-^{31}P/^{13}C. The ^{1}H and ^{13}C NMR chemical shift assignments of the polysaccharide material were carried out in D_2O solution (99.96%) at 323.2 K unless otherwise stated. Chemical shifts are reported in parts per million using internal sodium 3-trimethylsilyl-(2,2,3,3-$^2H_4$)-propanoate (TSP, δ_H 0.00 ppm), external 1,4-dioxane in D_2O (δ_C 67.40 ppm) and 2% H_3PO_4 in D_2O (δ_P 0.00 ppm) as reference. The ^{1}H,^{1}H-TOCSY experiments (dipsi2ph) were recorded with mixing times of 10, 30, 60, 90 and 120 ms. The ^{1}H,^{1}H-NOESY experiments[58] were collected with mixing times of 100 and 200 ms. Uniform and non-uniform sampling (50 and 25% NUS) were used for the multiplicity-edited ^{1}H,^{13}C-HSQC experiments[59] employing an echo/anti-echo-TPPI gradient selection with and without decoupling during the acquisition. The 2D ^{1}H,^{13}C-HSQC-TOCSY experiments were acquired using MLEV17 for homonuclear Hartman–Hahn mixing, an echo/anti-echo-TPPI gradient selection with decoupling during acquisition and mixing times of 20, 40, 80 and 120 ms. The 2D ^{1}H,^{31}P-Hetero-TOCSY experiments[60] were collected using a DIPSI2 sequence with mixing times of 10, 20, 30, 50 and 80 ms. The 2D ^{1}H,^{31}P-HMBC experiments were recorded using an echo/anti-echo gradient selection and mixing times of 25, 50 and 90 ms. The 3D ^{1}H,^{13}C,^{31}P[27] spectra were obtained using echo/anti-echo gradient selection and constant time in t_2 with a nominal value of $^nJ_{CP}$ of 5 Hz and without multiplicity selection. The spectra were processed and analyzed using TopSpin 4.0.1 software (Bruker BioSpin).

Statistical analysis. Unless otherwise indicated, statistical analysis was carried out on pooled data from at least three independent biological repeats. A two-way ANOVA with Bonferroni multiple comparison test was used to compare multiple groups. A P value equal to or less that 0.05 was considered statistically significant.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Illumina sequencing reads from the Tn-seq analysis were deposited in the NCBI Sequence Read Archive under the accession number SRP150081. The Tn-seq data, analyses and pipeline for the Tn-seq analyses are accessible under the DOI number 10.5281/zenodo.2541163 in GitHub as the following link: https://doi.org/10.5281/zenodo.2541163. Atomic coordinates and structure factors of the reported crystal structures have been deposited in the Protein Data Bank with accession codes 5U9Z (apo eGacH) and 6DGM (GroP●eGacH complex). All data generated during this study are included in the article, and supplementary information files or will be available from the corresponding authors upon reasonable request.

References

50. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. & Pease, L. R. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* **77**, 61–68 (1989).
51. Trevino, J., LiuZ., Cao, T. N., Ramirez-Pena, E. & Sumby, P. RivR is a negative regulator of virulencefactor expression in group A *Streptococcus*. *Infect. Immun.* **81**, 364–372 (2013).
52. Ghomashchi, F. et al. Preparation of the full set of recombinant mouse- and human-secreted phospholipases A2. *Methods Enzymol.* **583**, 35–69 (2017).
53. Le Breton, Y. & McIver, K. S. Genetic manipulation of *Streptococcus pyogenes* (the group A *Streptococcus*, GAS). *Curr. Protoc. Microbiol.* **30**, Unit 9D 3 (2013).
54. Kabsch, W. Xds. *Acta Crystallogr.* **66**, 125–132 (2010).
55. Adams, P. D. et al. *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr.* 66, 213–221 (2010).
56. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of *Coot*. *Acta Crystallogr.* **66**, 486–501 (2010).
57. Chen, V. B. et al. *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr.* **66**, 12–21 (2010).
58. Wagner, R. & Berger, S. Gradient-selected NOESY-A fourfold reduction of the measurement time for the NOESY Experiment. *J. Magn. Reson.* **123**, 119–121 (1996).
59. Willker, W., Leibfritz, D., Kerssebaum, R. & Bermel, W. Gradient selection in inverse heteronuclear correlation spectroscopy. *Magn. Reson. Chem.* **31**, 287–292 (1993).
60. Kellogg, G. W. Proton-detected hetero-TOCSY experiments with application to nucleic acids. *J. Magn. Reson.* **98**, 176–182 (1992).

nature research

Corresponding author(s): Nina van Sorge, Natalia Korotkova

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection

Macromolecular structural data were collected using Sergui (SER-CAT, APS, beamline 22-ID) and Blu-Ice 5 (SSRL, beamline 9-2). The data were processed using XDS and XSCALE, versions May 1, 2016 BUILT=20160617 and Jan 26, 2018, BUILT=20180126.
Mass-spectrometry data were collected using Thermo Xcalibur 4.0 (Thermo Fisher Scientific, Inc.).
The sequences of GacH, LtaS, and LtaP for construction of the phylogenetic tree were obtained from the ncbi database using blastp 2.7.1.
The NMR spectra were recorded on a Bruker AVANCE III 700 MHz equipped with a 5 mm TCI Z-Gradient Cryoprobe (1H/13C/15N) and dual receivers and a Bruker AVANCE II 600 MHz spectrometer equipped with a 5 mm TXI inverse Z-Gradient 1H/D-31P/13C. The spectra were acquired using TopSpin 3.1 software (Bruker BioSpin) for the 600 MHz machine and TopSpin 3.2 (Bruker BioSpin) for the 700 MHz machine.

Data analysis	Microsoft Excel 16.20 and 14.7.7, GraphPad Prism 5.0, SigmaPlot 14.0, PyMol v1.8.0.3, MolProbity 4.4, Grade Web Server 1.2.13, wwPDB Validation Service rb-20031172, Lipidsearch 4.1, TOPCONS 2.0, HHpred (May 19, 2017), TMHMM 2.0. Tn-seq data were analyzed using Bioperl: 1.7.1, Biopieces: 2.0, Cutadapt: 1.13 , DESeq2: 1.20.0, EdgeR: 3.22.3, hpgltools (version 2018.03), commit id: c730ef178f8e57bbf3819e21cf5e6cfe879e6328, limma: 3.36.2. Macromolecular structural data were analyzed using PHENIX and phenix.refine, versions dev_2481 and dev_3139; Phaser, version 2.8.2; Coot, version 0.8.9.1. Mass-spectrometry data were analyzed using Thermo Xcalibur 4.0 (Thermo Fisher Scientific, Inc.). Phylogenetic analysis was done using MEGA6 6.0 and MUSCLE 3.5. The NMR spectra were processed, analyzed and plotted using TopSpin 4.0.1 software (Bruker BioSpin).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The structure factors and coordinates were deposited to the Protein Data Bank and released with accession codes 5U9Z (apo eGacH) and 6DGM (GroP•eGacH complex).
Illumina sequencing reads from the Tn-seq analysis were deposited in the NCBI Sequence Read Archive (SRA) and released under the accession number SRP150081.
The Tn-seq data, analyses, and pipeline for the Tn-seq analyses are accessible under the DOI number 10.5281/zenodo.2541163 in GitHub as the following link http://doi.org/10.5281/zenodo.2541163.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Because for experiments performed in this manuscript, the sample size was not a factor of the analysis, no calculations of sample size were performed. All sample sizes were determined in accordance with published literature relevant to a particular experiment and they were optimal to generate statistically significant results. In general, the experimental analysis was performed in triplicate unless otherwise indicated. GC-MS measurements of carbohydrate composition and LC-MS analysis of GroP were performed twice. X-ray data were collected from single crystals.
Data exclusions	No data was excluded from analysis.
Replication	All replicates were performed in independent measurements and they were successful. Tn-seq data analysis including sample diagnostic analyses (PCA, heirarchical clustering) are accessible under the DOI number 10.5281/zenodo.2541163 in GitHub as the following link http://doi.org/10.5281/zenodo.2541163
Randomization	No randomization was applied to this study. Each biochemical assay was designed and performed with controls to avoid biased results. Crystal and glycopolymer samples are not required to be allocated into experimental groups in structural studies, and no animals or human research participants are involved in this study.
Blinding	No human participants or animals were employed in our works, so there was no issue about blinding for group allocation.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☐	☒ Unique biological materials
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Unique biological materials

Policy information about availability of materials

Obtaining unique materials: The bacterial strains and plasmids generated in this study are available upon request from the corresponding authors Natalia Korotkova and Nina van Sorge.

Antibodies

Antibodies used: Anti-Streptococcus pyogenes Group A Carbohydrate goat polyclonal antibodies, Abcam, ab9191, lots GR3231184-1 and GR282892-1; dilution 2,500.

Validation: The pAb ab9191 were validated previously by Rush et al., 2017 (JBC, 292, 19441-19457) and has been used extensively in GAS field. Further validation information is available on the vendor website: https://www.abcam.com/streptococcus-pyogenes-group-a-carbohydrate-antibody-ab9191.html