Molecular basis for recognition of the Group A Carbohydrate backbone by the PlyC streptococcal bacteriophage endolysin

Harley King[1,^], Sowmya Ajay Castro[2,^], Amol Arunrao Pohane[3], Cynthia M. Scholte[1], Vincent A. Fischetti[4], Natalia Korotkova[3,5], Daniel C. Nelson[1] and Helge C. Dorfmueller[2,*]

[1]Institute for Bioscience and Biotechnology Research, University of Maryland, Rockville, MD, USA

[2]Division of Molecular Microbiology, School of Life Sciences, University of Dundee, Dundee, UK

[3]Department of Microbiology, Immunology and Molecular Genetics, University of Kentucky, Lexington, Kentucky, USA

[4]Laboratory of Bacterial Pathogenesis and Immunology, The Rockefeller University, New York, NY, USA

[5]Department of Molecular and Cellular Biochemistry, University of Kentucky, Lexington, KY, USA

[^]equal contribution

[*]Correspondence: Helge C. Dorfmueller, h.c.z.dorfmueller@dundee.ac.uk

Keywords: Streptococcus; rhamnose, polysaccharide, cell wall, bacteriophage, endolysin, pyogenes

27 **Abbreviations (alphabetical order)**

28 Cell wall-binding domain (CBD), enzymatically active domain (EAD), Glucose (Glc), glycerol
29 phosphate (GroP), Glycosyl Hydrolase (GlyH), Group A Carbohydrate (GAC), Group A
30 Streptococcus (GAS), Group A-variant Streptococcus (GAVS), Group C Carbohydrate (GCC),
31 Group C Streptococcus (GCS), Group G Carbohydrate (GGC), *N*-acetyl-glucosamine (GlcNAc),
32 *N*-acetyl-galactosamine (GalNAc), N-acetyl-muramic acid (MurNAc), Peptidoglycan (PG),
33 polyrhamnose (pRha), Rhamnose (Rha), *S. dysgalactiae subsp. equisimilis* (SDSE), *S. mutans*
34 serotype c carbohydrate (SCC)

35

36

Abstract

Endolysins are peptidoglycan (PG) hydrolases that function as part of the bacteriophage (phage) lytic system to release progeny phage at the end of a replication cycle. Notably, endolysins alone can produce lysis without phage infection, which offers an attractive alternative to traditional antibiotics. Endolysins from phage that infect Gram-positive bacterial hosts contain at least one enzymatically active domain (EAD) responsible for hydrolysis of PG bonds and a cell wall binding domain (CBD) that binds a cell wall epitope, such as a surface carbohydrate, providing some degree of specificity for the endolysin. Whilst the EADs typically cluster into conserved mechanistic classes with well-defined active sites, relatively little is known about the nature of the CBDs and only a few binding epitopes for CBDs have been elucidated. The major cell wall components of many streptococci are the polysaccharides that contain the polyrhamnose (pRha) backbone modified with species-specific and serotype-specific glycosyl side chains. In this report, using molecular genetics, microscopy, flow cytometry and lytic activity assays, we demonstrate the interaction of PlyCB, the CBD subunit of the streptococcal PlyC endolysin, with the pRha backbone of the cell wall polysaccharides, Group A Carbohydrate (GAC) and serotype *c*-specific carbohydrate (SCC) expressed by the Group A Streptococcus and *Streptococcus mutans*, respectively.

56 **Introduction**

57 Endolysins are bacteriophage-encoded PG hydrolases that normally function from within the cell
58 to lyse the bacterial host, releasing progeny phage and completing the phage lifecycle [1].
59 However, the lytic activity of endolysins can be harnessed for antimicrobial use due to their ability
60 to equally lyse bacteria when applied exogenously, without infection by a parental phage. Due to
61 their direct lytic action on target PG, endolysins are not affected by efflux pumps, alterations in
62 metabolism, or other mechanisms of antibiotic resistance, making them ideal candidates for
63 development against multi-drug resistant organisms [2-4]. Notably, at least three endolysins,
64 some of which are active against methicillin-resistant *Staphylococcus aureus*, are currently being
65 evaluated in human clinical trials for their antimicrobial activity (reviewed in [5]).

66

67 Most endolysins, and in particular those from phage that infect Gram-positive bacterial hosts, are
68 comprised of modular domains. An enzymatically active domain (EAD) is generally found in the
69 N-terminal region, while a cell wall-binding domain (CBD) is located in the C-terminal region [6].
70 As the name implies, the EAD is a catalytic domain that is responsible for cleaving specific bonds
71 in the PG, the nature of which is dependent on the mechanistic class of the EAD. Occasionally,
72 endolysins contain two EADs, although both are not necessarily active. The CBD binds at high
73 affinity [7] to a cell wall-specific epitope and was suggested to dictate genus, species and serovar-
74 specificity of the endolysin. The CBD targets may be surface carbohydrates, wall teichoic acids
75 linked to the Gram-positive bacterial cell wall, or the PG itself [8].

76

77 The endolysin now known as PlyC is one of the first described endolysins and remains one of the
78 most studied. In 1934, Alice Evans noted a “nascent lysis” activity derived from streptococcal
79 phage lysates on streptococcal strains that were not sensitive to the phage itself [9]. By 1957,
80 Krause had determined that the phage used by Evans was specific for Group C Streptococci
81 (GCS), but an “enzyme” produced by the phage could lyse Groups A, A-variant, and C

82 Streptococci (GAS, GAVS, and GCS, respectively) [10]. These findings were confirmed by
83 Maxted later the same year and extended to include Group E Streptococci (GES) [11]. In 2001,
84 PlyC, then referred to as the streptococcal C_1 lysin, became the first endolysin to be used
85 therapeutically, protecting mice from GAS challenge in a nasopharyngeal model [2]. Subsequent
86 studies revealed that PlyC is a structurally unconventional endolysin, which is not encoded by a
87 single gene as found for all other endolysins described to date. Rather, PlyC is a nine-subunit
88 holoenzyme encoded by two distinct genes, *plyCB* and *plyCA*, within a polycistronic operon [12].
89 Eight PlyCB subunits self-assemble into a ring structure and form the basis of the CBD that binds
90 the streptococcal surface. A single PlyCA subunit contains two distinct EADs separated by an
91 extended α-helical linker region, which interfaces with the N-terminal residues of the PlyCB
92 octamer [13]. The PlyCA EADs consist of a glycosyl hydrolase (GH) domain and a cysteine,
93 histidine-dependent aminohydrolase/peptidase (CHAP) domain. The very high lytic activity of
94 PlyC relative to other endolysins is attributed to synergy/cooperativity between the two EADs.
95
96 Although the EADs of PlyCA have been extensively characterized with respect to specificity of
97 PG bonds they cleave, active-site residues, and their synergistic activity, it is unclear how PlyCB
98 recognizes PG of GAS, GCS and GES. Similar to other Gram-positive bacteria, the plasma
99 membrane of streptococci is surrounded by thick cell wall that consists of a complex network of
100 PG with covalently attached polysaccharides. Rebecca Lancefield utilized the unique
101 immunogenicity of the surface carbohydrates in β-hemolytic streptococci to subsequently
102 separate them into serogroups [14]. *S. pyogenes* is categorized as Group A Streptococcus, whilst
103 *S. dysgalactiae subsp equisimilis* (SDSE) produce at least two types of carbohydrates and are
104 annotated as Group C and G Streptococci, respectively [15]. Pioneering work by many
105 researchers have revealed that the cell wall polysaccharides of GAS, GCS and *S. mutans*, consist
106 of a pRha backbone modified with species-specific and serotype-specific glycosyl side chains
107 (Fig. 1A) [16, 17]. In GAS, GCS and *S. mutans* serotype *c*, the polysaccharides termed the

108 Lancefield group A carbohydrate (GAC), group C carbohydrate (GCC), and SCC, respectively,
109 have a conserved repeating →3)α-Rha(1→2)α-Rha(1→ di-saccharide backbone [16, 17]. The β-
110 *N*-acetylglucosamine (GlcNAc) side chains are attached to the 3-position of the α-1,2 linked
111 rhamnose (Rha) in GAC [16, 18, 19]. The GCC side chains have two *N*-acetyl-galactosamine
112 (GalNAc) residues attached to the same hydroxyl of Rha [16, 18, 20]. SCC carries the α-glucose
113 (Glc) side chains attached to the 2-position of the α-1,3 linked Rha in SCC [17]. Additionally, the
114 side chains of GAC and SCC are decorated in parts with the glycerol phosphate (GroP) moiety
115 [21]. The GAC and SCC biosynthetic pathways are encoded by the 12-gene loci
116 *gacABCDEFGHIJKL* and *sccABCDEFGHMNPQ*, respectively. The molecular mechanisms of
117 GAC and SCC biosynthesis have been the focus in a number of recent studies [21-25].
118
119 Remarkably, extracted GAC is known to partially inhibit the lytic actions of PlyC [26]. Furthermore,
120 a GAS mutant carrying the unmodified GAC (lacking the GlcNAc side chains) displayed enhanced
121 susceptibility to PlyC [25], implicating PlyCB in recognition of the pRha backbone. Here, we
122 identify the pRha backbone of GAC as the definite minimal binding epitope for PlyCB.
123
124 **Materials and Methods**
125 **Bacterial strains, growth conditions and media**
126 *Streptococcus pyogenes* strain D471 was propagated on solid media in plates containing Todd-
127 Hewitt broth supplemented with yeast extract (0.2%) (THY) and agar (1.4%) or in liquid THY as
128 described by Gera et al [27]. *S. mutans* wild type (Xc) and mutants were grown in Todd-Hewitt
129 broth with 1% yeast extract. All cultures were grown without antibiotics and without aeration at
130 37°C. *E. coli* genotypes DH5α (NEB, cat. No. C2988J), DH10α (Thermo Fisher, cat No.
131 18297010), BL21 (NEB, cat No. C2530H) and Origami 2 (Millipore Sigma, cat. No. 71346) were
132 used for routine plasmid propagation or protein expression and grown in Lysogeny Broth (LB)

133 medium supplemented with either 50 µg/ml kanamycin, 100 µg/ml ampicillin or 35 µg/ml
134 chloramphenicol as needed.
135
136 Bacterial strains *E. coli* CS2775 were transformed with pRGP1 plasmid [28], *gacABCDEFG* [23]
137 to produce pRha or empty plasmid control (pHD0131). The bacterial cells were grown overnight
138 in LB containing erythromycin (150 µg/ml) at 37°C and used next day for whole cell Western blots
139 and FACS and microscopy analysis.
140
141 **Recombinant expression and purification of $PlyCB_{WT}$ and $PlyCB_{R66E}$**
142 $PlyCB_{WT}$ (GenBank ID: NC_004814.1:7517-7735) was expressed from a pBAD24 vector and
143 purified from BL21 cells as previously described [12]. In brief, the culture was grown in LB and
144 induced with 0.25% L-arabinose at OD_{600} ~1.2-1.4 (Alfa Aesar, cat no. A11921). Cultures were
145 grown at 37°C with shaking at 180 RPM for 3-4 hours, centrifuged at 4500 x g and resuspended
146 in phosphate-buffered saline (PBS). Lysis was performed by a French press (1800 psi).
147 Benzonase (Millipore Sigma, cat. No. 70746-3) was added and the lysate incubated at room
148 temperature with rotation for 20-30 min. The lysate was centrifuged at 20,000 x g for 20 min and
149 the cleared lysate was passed through a 0.45-µm filter and loaded onto a XK-26/20 column
150 (Cytiva) with 30-35 ml ceramic hydroxyapatite (Bio-Rad, cat no. 1582000). $PlyCB_{WT}$ was eluted
151 from the column with three column volumes of 1 M sodium phosphate buffer (pH 7.2). Protein
152 was subsequently dialyzed in PBS, 10% glycerol and stored at -80°C until use. Protein purification
153 of $PlyCB_{R66E}$ was performed as $PlyCB_{WT}$ [13]. The fluorescent labeling of $PlyCB_{WT}$ and $PlyCB_{R66E}$
154 was performed using the manufacturer's recommended guidelines (Thermo Fisher, cat. No.
155 A20174). Analytical gel filtration was used to determine the multimeric nature of PlyCB and
156 $PlyCB_{R66E}$. A total of 100 µl (0.1 mg) of each protein was applied to a Superose 12 10/300 GL
157 column (Cytiva) and run in isocratic conditions in PBS buffer for 1.5 column volumes on an AKTA
158 FPLC system (Cytiva). Gel filtration standards (Bio-Rad) were also run under identical conditions.

159

160 **Purification of GAC**

161 GAS cells were grown overnight in THY media at 37˚C. Cultures were centrifuged at 4,500 x g.
162 Pellets were washed and resuspended in 40 ml distilled water per each original liter of media
163 used and combined in an 800 ml beaker. 22.5 ml 4N sodium nitrite (5 ml per liter of culture) was
164 added to the beaker in addition to 22.5 ml glacial acetic acid (5 ml per liter of culture). An orbital
165 shaker was used to gently mix the beaker for 15 min in a hood. The mixture was centrifuged in
166 500 ml bottles at 8000 x g for 15 min. The supernatant was decanted to a new beaker and
167 neutralized with 1M sodium hydroxide. The total solution, about 300 ml, was filtered with a 0.45-
168 micron filter assembly. 50-50 ml aliquots were deposited in a 3.5 kDa membrane and dialyzed in
169 a 4-liter beaker overnight with water. The following day, the solution was concentrated using an
170 Amicon 400 ml stirred cell (model 8400) filter assembly with a 76 mm diameter Ultracel® 5 kDa
171 ultrafiltration disc (Millipore Sigma, cat. No. PLCC07610) for 2 hours with 60 psi. The ~10 ml
172 concentrate was loaded onto an S-100 column for final purification. Fractions were assayed for
173 Rha, lyophilized and stored at 4°C until use.

174

175 **Calculation of pRha concentration**

176 A modified protocol by Edgar *et al.* [21] based on a protocol from DuBois *et al.* [29] was used to
177 determine Rha concentration in purified GAC. Briefly, anthrone reagent was prepared by
178 dissolving 0.2% w/w anthrone in H_2SO_4. Eighty microliters of aqueous samples or standards
179 containing either GAC or L-Rha at known concentrations were added to a 1.5 ml microfuge tube.
180 To this same tube, 320 µL of the anthrone reagent was added. Samples were boiled at 98°C for
181 10 min in a heat block. Samples were cooled to room temperature, transferred to a quartz plate,
182 and the absorbance at 580 nm was recorded using a spectrophotometer. Rha concentration was
183 interpolated using an L-Rha standard curve.

184

185 **Precipitation of PlyCB with GAC**

186 $PlyCB_{WT}$ and $PlyCB_{R66E}$ samples were defrosted from storage at -80°C. Lyophilized GAC was
187 resuspended in PBS. Both proteins and GAC were added to a 3.5 kDa dialysis membrane and
188 dialyzed overnight in PBS. Protein concentrations were determined using a NanoDrop
189 spectrophotometer (Thermo Fisher ND-2000) at 280 nm and were diluted to 5 mg/ml. The GAC
190 was also assayed and diluted with PBS to 1.6 mg/ml. One-hundred microliters of proteins and
191 100 µl of GAC or PBS were mixed in a 250 µl quartz plate and allowed to incubate without shaking
192 at room temperature. Visible precipitate formed in samples in 5-8 minutes. After recording the
193 precipitate at 340 nm using a Spectramax® M5 (Molecular Devices) spectrophotometer, the total
194 sample volume was transferred to a 1.5 ml microfuge tube. Samples were centrifuged at 14,000
195 x g to pellet the precipitate and supernatants were transferred to new 1.5 ml microfuge tubes.
196 Two-hundred microliters of 8 M urea was added to the pellet. Pellets were resuspended and 5 µl
197 of either pellet or supernatant were added to 40 µl water with 8 µl 6x Laemmli buffer with DTT.
198 Samples were boiled at 98°C for 8 min, and then 12.5 µl were loaded onto a 7.5% SDS-PAGE
199 and run for 32 min at 200 V. Proteins were visualized using Coomassie stain.

200

201 **Lysis assay**

202 A turbidity reduction assay was used to ascertain strain sensitivity to PlyC. This assay was
203 performed as previously described [30], except PlyC was used at 2 µM. Eight technical replicates
204 were performed.

205

206 Sensitivity of streptococcal species to PlyC-mediated lysis was analyzed using a wide range of
207 clinical isolates: 1) GAS isolates: $M1_{UK}$, WT 5448 strain [M1], delta*gacI* 5448 strain [d*gacI*]; 2)
208 GCS isolate: stC74A.0; 3) SDSE_*gac* isolates: stG245.0, stG652.0. stG485.0; GGS: stG6.0,
209 stG485.0 and 4) *S. mutans* serotype c were used as negative controls. Briefly, all streptococcal
210 strains were grown in THY at 37˚C overnight in 5% CO_2, except for *S. mutans*, which was grown

211 in THB media. Next day, the bacterial cells were inoculated in 1:100 fresh media and grown until
212 mid-logarithmic phase (OD_{600} 1.0). The cells were washed in PBS and resuspended to an OD_{600}
213 of 2.0. In a 96-well plate, to a 100 µl of bacterial cells, 100 µl of PlyC [1 µg/ml] was added and
214 immediately read at an absorbance of OD_{600}. The obtained values were standardized by
215 subtracting from the background values. The data is plotted using GraphPad Prism version 9.
216

217 **SDS-PAGE and blotting analysis**

218 $PlyCB_{WT}$ binding to recombinant *E. coli* expressing pRha was conducted using blot analysis.
219 Briefly, the lysate from the overnight cultures was analyzed in 20% tricine gels. SDS-PAGE and
220 protein transfers were performed according to manufactures instructions, Atto Ae-6050 Mini Gel
221 chamber and Novex protein separation from Thermo Fisher, respectively. The PVDF membranes
222 were blocked with 5% non-fat dry milk with Tris-Buffered Saline, 0.1% Tween® 20 detergent prior
223 to incubation with $PlyCB_{WT}$ labelled with Alexa Fluor® 647 (1:1000) for one hour at room
224 temperature. Goat anti-rabbit GAC antibodies conjugated with IRDye® 800CW were used as a
225 positive control (abcam ab216773). The resulting blots were imaged using the Licor Odyssey FC
226 Imaging System. All the blots were processed in parallel under the same conditions.
227

228 **Microscopy**

229 Microscopic analysis of *E. coli* bacteria was performed using cells that were grown overnight in
230 LB supplemented with antibiotics at 37˚C and diluted 1:100 the next day and regrown until OD_{600}
231 reached 0.5. The cells were washed twice with PBS for 5 min at 10,000 rpm and stained with
232 1:1000 dilution of $PlyCB^{AF647}$ and left for 20 minutes on ice in the darkness. Prior to fixing the cells
233 with 4% paraformaldehyde, the cells were washed again with PBS. The fixed cells were mounted
234 on 1% agarose coated microscopic slides and viewed under the CY5 channel on a fluorescent
235 Deltavision widefield microscope.
236

237 **FACS analysis**

238 *E. coli* cells were grown overnight as described above and diluted 1:100 the next day, grown at
239 37˚C and used for the assay at OD_{600} = 0.5. The cells were washed with PBS and probed with
240 $PlyCB^{AF647}$ or $PlyCB_{R66E}^{AF647}$ at 1:1000 dilution at 1:1000 dilution. Anti-GAC antibodies conjugated
241 with FITC (ABIN238144, antibodies-online, titer 1:50) were used as a positive control. The
242 samples were incubated for 20 minutes on ice at dark conditions. The cells were washed twice
243 with PBS at 14,000 rpm for 5 minutes and fixed with 4% paraformaldehyde. BD LSRFortessa
244 Flow Cytometry software was used to analyse the samples and the data interpretation was
245 conducted with FlowJo™ software v10.6.2.

246

247 **PlyC hydrolysis of sacculi**

248 *S. mutans* wild-type (WT) and the mutant strain sacculi were obtained by the SDS-boiling
249 procedure [25] followed by four washes each with 1 M NaCl and distilled water. The sacculi were
250 resuspended in PBS to OD_{600} of 1.0 and incubated with PlyC (5 µg/ml) in a 96-well plate. The
251 lysis was monitored after 10, 20, 30, 40, 50 and 60 min as a decrease in OD_{600}. Results were
252 reported as fold change in OD_{600} of the sacculi incubated with PlyC vs the sacculi incubated
253 without PlyC.

254

255

256 **Results & Discussion**

257

258 **Pathogenic streptococci producing the GAC are susceptible to PlyC**

259 A major component of the GAS cell wall is the GAC, building approximately 50% of the cell wall
260 by weight [31]. The GAC is universally conserved amongst all isolated GAS strains on the basis
261 of the gene cluster sequence conservation [32]. A number of *S. dysgalactiae subsp. equisimilis*
262 (SDSE) isolates, naturally belonging to Group G Streptococci (GGS), have been reported to have

263 undergone homologous recombination and replaced their Group G Carbohydrate (GGC) gene
264 cluster in parts with the GAC gene cluster [33-35]. We therefore expanded the previously reported
265 PlyC streptococci cell lysis assay used by Nelson *et al.* [2] to investigate those new isolates named
266 SDSE_*gac*. We also tested if PlyC was able to lyse a selection of GAS serotypes including a
267 newly emerged isolate $M1_{UK}$ [36], and included negative controls GGS isolates and *S. mutans*
268 serotype C (Fig. 1B). In agreement with the published literature, all tested GAS serotypes are
269 susceptible to PlyC lysis and the two GGS isolates are resistant to PlyC. The GGC does not
270 contain the GAC, GCC and SCC pRha backbone with →3)α-Rha(1→2)α-Rha(1→ di-saccharide
271 repeats, but an alternating Rha-Gal carbohydrate [37] (Fig. 1A). All three SDSE_*gac* isolates
272 tested have inherited parts of the *gac* gene cluster and produce the GAC instead of the GGC [34].
273 Strikingly, the SDSE_*gac* strains are all sensitive to PlyC treatment. The fact that PlyC is able to
274 lyse SDSE_*gac* strains that express the GAC, but PlyC does not lyse the isogenic GGS strains
275 producing the GGC strongly suggests that the GAC is a critical component of PlyC recognition
276 and subsequent activity.
277
278 Importantly, all strains tested in this study that are susceptible to PlyC lysis produce a cell wall
279 polysaccharide that contains the pRha backbone and a β-linked sugar substituent on the α1,2-
280 linked Rha (Fig 1A). We therefore suggest that the pRha backbone with and without a side chain
281 are both vital ligands to assist PlyC activity and the new SDSE_*gac* isolates will also be
282 susceptible to PlyC treatment due to production of the GAC.
283
284 **Purified GAC precipitates PlyCB - but not $PlyCB_{R66E}$**
285 The lysis assay of GAS cells, and in particular of the SDSE_*gac* variants, suggests that either the
286 ubiquitous pRha or GAC in GAS cells is the ligand for PlyCB. We propose that the PlyCB
287 octameric CBD binds GAC and/or the GAC pRha backbone. We tested this hypothesis by
288 investigating the binding of PlyC to partially purified GAC. We hypothesized if the GAC was able

289 to precipitate PlyCB, an interaction of the two systems must have occurred [38]. As a negative
290 control, we employed the previously published inactive mutant $PlyCB_{R66E}$, which lost the ability to
291 bind to GAS cells [13]. Consistent with prior analytical gel filtration and dynamic light scattering
292 results [12], as well as the crystal structures [13, 39, 40], both PlyCB and its R66E mutant self-
293 assemble into stable octameric structures (Fig 1C, SF 1). The purified proteins were incubated
294 with the extracted GAC, and precipitation was monitored at 340 nm, a standard wavelength for
295 measuring protein aggregation [41, 42] (Fig. 2A, B). Whilst keeping the PlyCB concentration
296 constant, we varied the concentration of GAC. Within five minutes at room temperature the
297 solution became turbid, suggesting aggregation (Fig. 2A). When the PlyCB concentration was
298 kept constant and the GAC concentration was varied, the turbidity correlated with PlyCB
299 concentration in a dose dependent manner, suggesting that PlyCB requires GAC to aggregate.
300 Importantly, PlyCB did not self-aggregate when no GAC was added in the assay. Furthermore,
301 no aggregation was detected when $PlyCB_{R66E}$ was incubated with purified GAC (Fig. 2B). To
302 demonstrate the presence of PlyCB in the precipitates, we analyzed the soluble and pellet
303 fractions (Fig. 2C, D). A higher yield of aggregated PlyCB was found in the pelleted samples when
304 compared to the soluble fraction (Fig. 2D). A similar precipitation effect was observed when we
305 varied the PlyCB concentration and kept the GAC concentration constant (Fig. 2E, F),
306 demonstrating that both species are necessary for an interaction.
307
308
309 **PlyCB binds to recombinantly produced pRha backbone**
310 The purified GAC from bacteria contains a mixture of carbohydrates, including the fully decorated
311 GAC with GroP[21] and a small proportion of the polysaccharide lacking the side chains [43]. PlyC
312 is able to lyse a number of GAS mutants including GAVS and d*gacI*_M1 [2, 11, 25, 44], that
313 decorate the cell wall with the unmodified GAC lacking the side chains (Fig. 1A, B), suggesting
314 that the pRha backbone of the GAC is the minimal carbohydrate structure required for PlyCB

315 binding. To test this hypothesis, we recombinantly produced the pRha backbone in *E. coli* cells.
316 We and others have previously reported that the *S. mutans sccABCDEFG* gene cluster, when
317 transformed into *E. coli* cells, functionally produces the pRha backbone attached to the lipid A
318 [23, 28]. Additionally, to understand if PlyCB recognizes a specific pRha backbone, we
319 engineered *E. coli* cells expressing the GAC gene cluster *gacABCDEFG* required for the GAC
320 pRha production. *E. coli* cells carrying an empty plasmid were used as a negative control.
321
322 Next, we investigated the binding of PlyCB conjugated with Alexa Fluor® 647 ($PlyCB^{AF647}$) to an
323 *E. coli* total cell lysate expressing the pRha backbone of the SCC or GAC, respectively (Fig. 3A).
324 The blotted membranes were incubated with $PlyCB^{AF647}$, and a positive interaction between
325 $PlyCB^{AF647}$ and the *E. coli* produced pRha is visualized as a prominent band around 40 kDa. The
326 size of the band agrees with the band detected by anti-GAC antibodies that were previously
327 reported to recognize pRha [23] (Fig. 3B). Importantly, $PlyCB^{AF647}$ and GAC antibodies do not
328 interact with the cell lysate of *E. coli* expressing an empty plasmid (Fig. 3 A, B). Contrary, the
329 R66E-mutant protein was not able to detect the pRha of the PAGE separated sample (SF 2). We
330 further confirmed the ability of PlyCB to bind to *E. coli* cells decorated with the pRha by fluorescent
331 microscopy (SF 3). Only cells that produce the pRha are detected by the $PlyCB^{AF647}$, in agreement
332 with the results of the blot analysis.
333
334 To gather additional evidence that PlyCB interacts with the pRha backbone, we established a flow
335 cytometry assay to analyse the binding of $PlyCB^{AF647}$ to pRha-producing *E. coli*. Unstained *E. coli*
336 cells that express pRha or carry an empty plasmid sort in the identical range (Fig. 4A). The GAC
337 antibodies label exclusively the cells producing pRha (Fig. 4A). A similar pattern of the GAC
338 antibodies binding is observed when the cells were incubated with $PlyCB^{AF647}$ (Fig. 4B). Contrary,
339 the $PlyCB_{R66E}{}^{AF647}$ mutant protein does not bind to *E. coli* cells, and $PlyCB^{AF647}$ does not interact
340 with the cells expressing an empty vector (Fig. 4B). Taken together, these data provide the first

341 definitive evidence that the pRha backbone of GAC and SCC is a binding receptor of the PlyCB
342 octameric subunit.
343
344 **PlyC lyses engineered *S. mutans* producing the GAC**
345 Despite the fact that the SCC pRha backbone is identical to the GAC, *S. mutans* is resistant to
346 PlyC lysis (Fig. 1A). To get a better understanding why *S. mutans* is resistant to PlyC, we
347 compared PlyC-induced lysis of the sacculi purified from *S. mutans* WT and a number of mutant
348 strains producing different SCC variants (Fig. 5). First, we examined the Δ*sccH* mutant producing
349 the GroP-deficient SCC [21]. Similar to *S. mutans* WT, Δ*sccH* was resistant to PlyC-mediated
350 lysis (Fig. 5). Second, we tested the *sccN* deletion mutant, Δ*sccN*, that is deficient in the enzyme
351 required for generation of the Glc side chains [45]. A time dependent lysis is observed for Δ*sccN*.
352 Expression of the WT copy of *sccN* in Δ*sccN* (the Δ*sccN*:p*sccN* strain) fully restored the resistance
353 of the bacteria to PlyC (Fig. 5). These observations clearly suggest that PlyC is able to bind to the
354 *S. mutans* cells producing the unmodified pRha backbone, and the Glc side chains in SCC hinders
355 PlyC binding. We then investigated whether the addition of the GAC GlcNAc side chains to the
356 pRha backbone affects sensitivity of the engineered *S. mutans* sacculi to PlyC-induced lysis. We
357 expressed the GAS genes *gacHIJKL* required for the formation and addition of the GlcNAc side
358 chains and GroP to GAC [21], in the Δ*sccN* background strain in two versions, creating the
359 Δ*sccN:pgacHI*JKL* and Δ*sccN:pgacHIJKL* strains [45]. The plasmid *pgacHI*JKL* contains an
360 inserted stop codon in the *gacI* gene required for generation of the GlcNAc side chain, and,
361 therefore, the Δ*sccN:pgacHI*JKL* strain produces the unmodified SCC lacking any side chains
362 (Fig. 5). As expected, the sacculi isolated from this strain remains susceptible to PlyC lysis. We
363 previously showed that in Δ*sccN:pgacHIJKL*, the Glc side chains are replaced with the GlcNAc
364 side chains [45]. Interestingly, expression of *gacHIJKL* in Δ*sccN* did not restore the resistance of
365 the bacteria to PlyC (Fig. 5), indicating that the GlcNAc side chains do not obstruct PlyC binding.
366 Lastly, we analyzed PlyC-mediated lysis of sacculi purified from the Δ*rgpG* mutant, which is

367 deficient in SCC expression [45]. The RgpG protein catalyzes the first step in SCC biosynthesis
368 [46]. In comparison to Δ*sccN*, PlyC-induced lysis of Δ*rgpG* was less pronounced (Fig. 5),
369 indicating the importance of the pRha backbone of SCC in PlyC activity and supporting the
370 findings that the pRha backbone is a ligand contributing to PlyC binding. These studies reveal
371 that if the SCC is 'unmasked' *i.e.,* stripped of the Glc and Glc-GroP side chains, it becomes a
372 ligand for PlyCB and that *S. mutans* is PlyC susceptible if SCC is replaced with GAC.
373
374
375 **Concluding remarks**
376
377 A structural feature of the PlyCB protein remains to be discovered that explains why only certain
378 streptococci are susceptible to PlyC's lytic activity. The pRha decorated with an α–linked side
379 chain sugar appears not compatible with the PlyCB ligand binding site and therefore only those
380 streptococci expressing pRha decorated with β-linked substituents, such as GlcNAc and GlcNAc-
381 GroP are susceptible. This could potentially be exploited for diagnostic purposes or in the case of
382 SCC and Group G Streptococci, opens up the potential for novel therapeutic approaches. If SCC
383 was treated by PlyC in combination with an additional enzyme that removes the α–linked side
384 chains in these streptococcal carbohydrates, this would expose the pRha backbone and
385 subsequently make these strains susceptible. A recently accepted manuscript by Boendum *et al.*
386 [40] reported 19 potential binding states of tetrarhamnose to PlyCB. Once the pRha ligand site
387 has been identified, this could also be further exploited by directed evolution approaches to
388 generate PlyCB protein variants that are capable to bind the carbohydrates from, for example,
389 SCC and Group G Streptococci.
390
391 Whilst much has been learned about the structure and function of PlyC in the past 20 years, many
392 questions remain, specifically with respect to its interaction with the PG. Considering the average

393 length of the cellular pRha is 7-10 kDa [25] and that the α1,2-1,3-pRha with or without β-
394 configured GlcNAc/GalNAc-side chains bind to PlyCB, it is inviting to speculate that an element
395 of avidity is responsible for tight binding of the PlyCB octamer to the streptococcal surface, in
396 agreement with the recent accepted manuscript [40]. Further proof is needed to substantiate this
397 hypothesis. Another question lies in the actions of the EADs relative to the PlyCB octamer. PlyC
398 clearly has a high turnover as demonstrated in multiple biochemical assays. However, it is
399 unknown if PlyCB "docks" to the surface and the flexibility of the holoenzyme allows the PlyC
400 EADs to cleave multiple bonds in a localized area weakening the overall superstructure of the
401 PG. Alternatively, the enzymatic turnover could be dictated by a balance of on and off rates of the
402 EADs and CBD monomers leading to widespread hydrolysis of the PG. Lastly, it is unknown
403 whether PlyC binds, cleaves, and releases the PG at random points on the streptococcal surface
404 or works its way down a single strand of PG in a processive manner. It is noteworthy that cellulase
405 enzymes, which cleave the β1,4 glycosidic linkages in cellulose, possess a catalytic domain, a
406 flexible linker, and a cellulose binding domain, analogous to the traditional endolysins. It has been
407 demonstrated that energy is stored in the flexible linker can adopt compact and extended
408 configurations that allows the cellulase to move in a "caterpillar-like" motion down a chain of
409 cellulose [47, 48]. Although PlyC does not contain an equivalent flexible linker, the octameric
410 nature of PlyCB invites the possibility that it may interact with the successive pRha strands
411 allowing PlyC to depolymerize the PG in a zipper-like fashion.
412
413 In conclusion, the α1,2-1,3-pRha is the definite, minimal carbohydrate substrate for the PlyCB
414 subunit. We validated this by comprehensive experiments, using pRha recombinantly produced
415 in *E. coli*, and the *S. mutans* variants of the Rha-based polysaccharides. The work described here
416 provides valuable insight into the molecular interactions that define a PlyC's host specificity, which
417 can inform the future studies as well as engineering approaches.

418

419 **Data Availability Statement**

420 No mandated datasets are associated with the paper.

421

422

423

424

425

426

427

428

429 **Figure Legend**

430 Fig 1)

431 A) Symbolic drawings of the carbohydrate structures of GAS (GAC), polyrhamnose backbone
432 (pRha) in GAVS and M1_d*gacI*, GCS (GCC), GGS (GGC) and *S. mutans* (SCC). For simplicity,
433 the reported glycerol phosphate of occasionally present on the GAC and SCC side chains have
434 been omitted. Repeat units are marked with brackets. The pRha backbone with alternating (α1-
435 >2) and (α1->3) linkages, whilst the α1->2 Rha being decorated with a β1->3 side chain is a
436 commonality among all PlyC susceptible strains. B) PlyC lysis of streptococcal pathogens. Group
437 A and Group C Streptococci serotypes are susceptible to PlyC-mediated lysis. Group G
438 Streptococci show limited susceptibility and *S. mutans* is resistant to PlyC lysis. Three SDSE
439 isolates that produce the GAC instead of GGC (SDSE_*gac*) are susceptible to PlyC treatment. C)
440 Bottom and top view onto WT PlyCB octameric structure with the GlyH and EAD domains omitted
441 (4f78.pdb).

442

443 Fig 2)

444 Precipitation studies of purified PlyCB and GAC reveal direct interaction of PlyCB with GAC. A)
445 The PlyCB concentration is kept constant whilst the GAC concentration is varied. Visible
446 precipitate forms at the higher concentrations. B) The precipitate level is measured
447 spectrophotometrically at 340 nm and compared to the mutant $PlyCB_{R66E}$, which does not bind
448 the GAC. C) Coomassie stained and D) densitometry analysis of PlyCB protein from the
449 supernatant fraction and aggregates (pellets). E, F) The same dose dependency is observed
450 when the PlyCB concentration is varied. Arrowhead depicts PlyCB protein at 8 kDa.
451
452 Fig 3)
453 Representation of immunoblot analysis of the cell lysate of *E. coli* expressing the SCC and GAC
454 pRha and carrying an empty control plasmid (-ve). A) Blot was incubated with PlyCB-Alexa Fluor®
455 647 ($PlyCB_{WT}^{AF647}$). B) Probing the same samples with the GAC antibodies confirms the presence
456 of GAC in the bands. Molecular mass markers are given in kDa.
457
458 Fig 4)
459 PlyCB binding to *E. coli* cells were investigated by flow cytometry after labelling with $PlyCB_{WT}^{AF647}$
460 and $PlyCB_{R66E}^{AF647}$ mutant proteins. Blue: -ve control cells without pRha. Red: pRha producing *E.*
461 *coli* cells. Representative histograms are shown. A) Left panel: unstained cells. Right panel: The
462 anti-GAC antibodies (GAC-FITC) were used as a positive control to label the *E. coli* cells
463 producing pRha. The antibodies do not bind to the *E. coli* cells carrying an empty plasmid (-ve).
464 B) Left panel: $PlyCB_{WT}^{AF647}$ binds to the *E. coli* cells producing pRha, but not to the *E. coli* cells
465 carrying an empty plasmid (-ve). Right panel: $PlyCB_{R66E}^{AF647}$ does not binds to the *E. coli* cells
466 producing pRha.
467
468 Fig 5)

469 The PlyC-mediated lysis of sacculi purified from *S. mutans* strains. The lysis was monitored after
470 10, 20, 30, 40, 50 and 60 min as a decrease in OD_{600}. Results are presented as a fold change in
471 OD_{600} of the sacculi incubated with PlyC *vs.* the sacculi incubated without PlyC. Data points and
472 error bars represent mean values of four biological replicates and standard deviation,
473 respectively. P-values were determined by two-way ANOVA with Dunnett's multiple comparisons
474 test.
475
476
477 **FUNDING**
478 The HCD laboratory is supported by Wellcome and Royal Society Grant 109357/Z/15/Z and the
479 University of Dundee Wellcome Trust Funds 105606/Z/14/Z and Tenovus Scotland Large
480 Research Grant [T17/17]. HK is supported by the National Institute of Standards and Technology
481 (NIST). VAF is supported by Rockefeller University laboratory funds. NK is supported by NIH
482 grants R01 DE028916 from the NIDCR and R01 AI143690 from the NIAID.
483
484 **ACKNOWLEDGEMENTS**
485 We thank the CDC Streptococcus Lab and the Active Bacterial Core surveillance program for
486 sharing isolates isolate 20170556 (stG6.0), 20171682 (stC74A.0), 20154376 (stG245.0),
487 20173686 (stG652.0), 20170560 and 20176966 (stG485.0). We also thank Ryan Heselpoth, Yang
488 Shen, and Sara Linden for reagents, technical advice, and helpful discussion.
489
490
491 **References**
492 1 Fernandes, S. and Sao-Jose, C. (2018) Enzymes and Mechanisms Employed by Tailed
493 Bacteriophages to Breach the Bacterial Cell Barriers. Viruses. **10**

494 2 Nelson, D., Loomis, L. and Fischetti, V. A. (2001) Prevention and elimination of upper
495 respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme.
496 Proc Natl Acad Sci U S A. **98**, 4107-4112
497 3 Dams, D. and Briers, Y. (2019) Enzybiotics: Enzyme-Based Antibacterials as
498 Therapeutics. Adv Exp Med Biol. **1148**, 233-253
499 4 Love, M. J., Bhandari, D., Dobson, R. C. J. and Billington, C. (2018) Potential for
500 Bacteriophage Endolysins to Supplement or Replace Antibiotics in Food Production and Clinical
501 Care. Antibiotics (Basel). **7**
502 5 Schmelcher, M. and Loessner, M. J. (2020) Bacteriophage endolysins - extending their
503 application to tissues and the bloodstream. Curr Opin Biotechnol. **68**, 51-59
504 6 Heselpoth, R. D., Swift, S. M., Linden, S. B., Mitchell, M. S. and Nelson, D. C. (2018)
505 Enzybiotics: endolysins and bacteriocins. In Bacteriophages - biology, technology, and therapy
506 (Harper, D., Abedon, S. T., Burrowes, B. and McConvill, M., eds.), Springer International
507 Publishing, Dordrecht, The Netherlands
508 7 Loessner, M. J., Kramer, K., Ebel, F. and Scherer, S. (2002) C-terminal domains of
509 Listeria monocytogenes bacteriophage murein hydrolases determine specific recognition and
510 high-affinity binding to bacterial cell wall carbohydrates. Molecular microbiology. **44**, 335-349
511 8 Mesnage, S., Dellarole, M., Baxter, N. J., Rouget, J. B., Dimitrov, J. D., Wang, N.,
512 Fujimoto, Y., Hounslow, A. M., Lacroix-Desmazes, S., Fukase, K., Foster, S. J. and Williamson,
513 M. P. (2014) Molecular basis for bacterial peptidoglycan recognition by LysM domains. Nat
514 Commun. **5**, 4269
515 9 Evans, A. C. (1934) *Streptococcus* bacteriphage: a study of four serological types.
516 Public Health Rep. **49**, 1386-1401
517 10 Krause, R. M. (1957) Studies on bacteriophages of hemolytic streptococci. I. Factors
518 influencing the interaction of phage and susceptible host cell. J Exp Med. **106**, 365-384
519 11 Maxted, W. R. (1957) The active agent in nascent phage lysis of streptococci. J Gen
520 Microbiol. **16**, 584-595
521 12 Nelson, D., Schuch, R., Chahales, P., Zhu, S. and Fischetti, V. A. (2006) PlyC: a
522 multimeric bacteriophage lysin. Proc Natl Acad Sci U S A. **103**, 10765-10770
523 13 McGowan, S., Buckle, A. M., Mitchell, M. S., Hoopes, J. T., Gallagher, D. T., Heselpoth,
524 R. D., Shen, Y., Reboul, C. F., Law, R. H., Fischetti, V. A., Whisstock, J. C. and Nelson, D. C.
525 (2012) X-ray crystal structure of the streptococcal specific phage lysin PlyC. Proc Natl Acad Sci
526 U S A. **109**, 12752-12757
527 14 Lancefield, R. C. (1933) A Serological Differentiation of Human and Other Groups of
528 Hemolytic Streptococci. J Exp Med. **57**, 571-595
529 15 Vandamme, P., Pot, B., Falsen, E., Kersters, K. and Devriese, L. A. (1996) Taxonomic
530 study of lancefield streptococcal groups C, G, and L (Streptococcus dysgalactiae) and proposal
531 of S. dysgalactiae subsp. equisimilis subsp. nov. Int J Syst Bacteriol. **46**, 774-781
532 16 Coligan, J. E., Kindt, T. J. and Krause, R. M. (1978) Structure of the streptococcal
533 groups A, A-variant and C carbohydrates. Immunochemistry. **15**, 755-760
534 17 Pritchard D G, G. R. L., Michlek S M, McGhee J R. . (1986) Biochemical aspects of
535 serotype carbohydrate antigens of Streptococcus mutans. In: Hamada S, Michalek S M, Kiyono
536 H, Menaker L, McGhee J R, editors. Molecular microbiology and immunology of Streptococcus
537 mutans. Amsterdam, The Netherlands: . Elsevier Science Publishers; , pp. 39–49.
538 18 Coligan, J. E., Schnute, W. C., Jr. and Kindt, T. J. (1975) Immunochemical and chemical
539 studies on streptococcal group-specific carbohydrates. J Immunol. **114**, 1654-1658
540 19 Huang, D. H., Rama Krishna, N. and Pritchard, D. G. (1986) Characterization of the
541 group A streptococcal polysaccharide by two-dimensional ^{1}H-nuclear-magnetic-resonance
542 spectroscopy. Carbohydrate research. **155**, 193-199

20 Neiwert, O., Holst, O. and Duda, K. A. (2014) Structural investigation of rhamnose-rich polysaccharides from Streptococcus dysgalactiae bovine mastitis isolate. Carbohydr Res. **389**, 192-195

21 Edgar, R. J., van Hensbergen, V. P., Ruda, A., Turner, A. G., Deng, P., Le Breton, Y., El-Sayed, N. M., Belew, A. T., McIver, K. S., McEwan, A. G., Morris, A. J., Lambeau, G., Walker, M. J., Rush, J. S., Korotkov, K. V., Widmalm, G., van Sorge, N. M. and Korotkova, N. (2019) Discovery of glycerol phosphate modification on streptococcal rhamnose polysaccharides. Nature chemical biology. **15**, 463-471

22 van der Beek, S. L., Zorzoli, A., Canak, E., Chapman, R. N., Lucas, K., Meyer, B. H., Evangelopoulos, D., de Carvalho, L. P. S., Boons, G. J., Dorfmueller, H. C. and van Sorge, N. M. (2019) Streptococcal dTDP-L-rhamnose biosynthesis enzymes: functional characterization and lead compound identification. Molecular microbiology. **111**, 951-964

23 Zorzoli, A., Meyer, B. H., Adair, E., Torgov, V. I., Veselovsky, V. V., Danilov, L. L., Uhrin, D. and Dorfmueller, H. C. (2019) Group A, B, C, and G Streptococcus Lancefield antigen biosynthesis is initiated by a conserved alpha-d-GlcNAc-beta-1,4-l-rhamnosyltransferase. J Biol Chem. **294**, 15237-15256

24 van der Beek, S. L., Le Breton, Y., Ferenbach, A. T., Chapman, R. N., van Aalten, D. M., Navratilova, I., Boons, G. J., McIver, K. S., van Sorge, N. M. and Dorfmueller, H. C. (2015) GacA is essential for Group A Streptococcus and defines a new class of monomeric dTDP-4-dehydrorhamnose reductases (RmlD). Molecular microbiology. **98**, 946-962

25 Rush, J. S., Edgar, R. J., Deng, P., Chen, J., Zhu, H., van Sorge, N. M., Morris, A. J., Korotkov, K. V. and Korotkova, N. (2017) The molecular mechanism of N-acetylglucosamine side-chain attachment to the Lancefield group A carbohydrate in Streptococcus pyogenes. J Biol Chem. **292**, 19441-19457

26 Lood, R., Raz, A., Molina, H., Euler, C. W. and Fischetti, V. A. (2014) A highly active and negatively charged Streptococcus pyogenes lysin with a rare D-alanyl-L-alanine endopeptidase activity protects mice against streptococcal bacteremia. Antimicrob Agents Chemother. **58**, 3073-3084

27 Gera, K. and McIver, K. S. (2013) Laboratory growth and maintenance of Streptococcus pyogenes (the Group A Streptococcus, GAS). Curr Protoc Microbiol. **30**, 9D 2 1-9D 2 13

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

29 DuBois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A. and Smith, F. (1956) Colorimetric method for determination of sugars and related substances. Anal. Chem. **28**, 350-356

30 Heselpoth, R. D., Yin, Y., Moult, J. and Nelson, D. C. (2015) Increasing the stability of the bacteriophage endolysin PlyC using rationale-based FoldX computational modeling. Protein Eng Des Sel. **28**, 85-92

31 M., 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

32 Davies, M. R., McIntyre, L., Mutreja, A., Lacey, J. A., Lees, J. A., Towers, R. J., Duchene, S., Smeesters, P. R., Frost, H. R., Price, D. J., Holden, M. T. G., David, S., Giffard, P. M., Worthing, K. A., Seale, A. C., Berkley, J. A., Harris, S. R., Rivera-Hernandez, T., Berking, O., Cork, A. J., Torres, R., Lithgow, T., Strugnell, R. A., Bergmann, R., Nitsche-Schmitz, P., Chhatwal, G. S., Bentley, S. D., Fraser, J. D., Moreland, N. J., Carapetis, J. R., Steer, A. C., Parkhill, J., Saul, A., Williamson, D. A., Currie, B. J., Tong, S. Y. C., Dougan, G. and Walker, M. J. (2019) Author Correction: Atlas of group A streptococcal vaccine candidates compiled using large-scale comparative genomics. Nat Genet. **51**, 1295

592 33 Brandt, C. M., Haase, G., Schnitzler, N., Zbinden, R. and Lutticken, R. (1999)
593 Characterization of blood culture isolates of Streptococcus dysgalactiae subsp. equisimilis
594 possessing Lancefield's group A antigen. J Clin Microbiol. **37**, 4194-4197
595 34 Chochua, S., Metcalf, B. J., Li, Z., Rivers, J., Mathis, S., Jackson, D., Gertz, R. E., Jr.,
596 Srinivasan, V., Lynfield, R., Van Beneden, C., McGee, L. and Beall, B. (2017) Population and
597 Whole Genome Sequence Based Characterization of Invasive Group A Streptococci Recovered
598 in the United States during 2015. mBio. **8**
599 35 McMillan, D. J., Vu, T., Bramhachari, P. V., Kaul, S. Y., Bouvet, A., Shaila, M. S.,
600 Karmarkar, M. G. and Sriprakash, K. S. (2010) Molecular markers for discriminating
601 Streptococcus pyogenes and S. dysgalactiae subspecies equisimilis. Eur J Clin Microbiol Infect
602 Dis. **29**, 585-589
603 36 Lynskey, N. N., Jauneikaite, E., Li, H. K., Zhi, X., Turner, C. E., Mosavie, M., Pearson,
604 M., Asai, M., Lobkowicz, L., Chow, J. Y., Parkhill, J., Lamagni, T., Chalker, V. J. and
605 Sriskandan, S. (2019) Emergence of dominant toxigenic M1T1 Streptococcus pyogenes clone
606 during increased scarlet fever activity in England: a population-based molecular epidemiological
607 study. Lancet Infect Dis. **19**, 1209-1218
608 37 Pritchard, D. G., Rener, B. P., Furner, R. L., Huang, D. H. and Krishna, N. R. (1988)
609 Structure of the group G streptococcal polysaccharide. Carbohydr Res. **173**, 255-262
610 38 Mirica, K. A., Lockett, M. R., Snyder, P. W., Shapiro, N. D., Mack, E. T., Nam, S. and
611 Whitesides, G. M. (2012) Selective precipitation and purification of monovalent proteins using
612 oligovalent ligands and ammonium sulfate. Bioconjug Chem. **23**, 293-299
613 39 Shen, Y., Barros, M., Vennemann, T., Gallagher, D. T., Yin, Y., Linden, S. B., Heselpoth,
614 R. D., Spencer, D. J., Donovan, D. M., Moult, J., Fischetti, V. A., Heinrich, F., Losche, M. and
615 Nelson, D. C. (2016) A bacteriophage endolysin that eliminates intracellular streptococci. Elife. **5**
616 40 Broendum, S. S., Williams, D. E., Hayes, B. K., Kraus, F., Fodor, J., Clifton, B. E., Geert
617 Volbeda, A., Codee, J. D. C., Riley, B. T., Drinkwater, N., Farrow, K. A., Tsyganov, K.,
618 Heselpoth, R. D., Nelson, D. C., Jackson, C. J., Buckle, A. M. and McGowan, S. (2021) High
619 avidity drives the interaction between the streptococcal C1 phage endolysin, PlyC, with the cell
620 surface carbohydrates of Group A Streptococcus. Molecular microbiology
621 41 Sato, T., Mattison, K. W., Dubin, P. L., Kamachi, M. and Morishima, Y. (1998) Effect of
622 protein aggregation on the binding of lysozyme to pyrene-labeled polyanions. Langmuir. **14**,
623 5430-5437
624 42 Wilcken, R., Wang, G., Boeckler, F. M. and Fersht, A. R. (2012) Kinetic mechanism of
625 p53 oncogenic mutant aggregation and its inhibition. Proc Natl Acad Sci U S A. **109**, 13584-
626 13589
627 43 Michon, F., Moore, S. L., Kim, J., Blake, M. S., Auzanneau, F. I., Johnston, B. D.,
628 Johnson, M. A. and Pinto, B. M. (2005) Doubly branched hexasaccharide epitope on the cell
629 wall polysaccharide of group A streptococci recognized by human and rabbit antisera. Infect
630 Immun. **73**, 6383-6389
631 44 van Sorge, N. M., Cole, J. N., Kuipers, K., Henningham, A., Aziz, R. K., Kasirer-Friede,
632 A., Lin, L., Berends, E. T. M., Davies, M. R., Dougan, G., Zhang, F., Dahesh, S., Shaw, L., Gin,
633 J., Cunningham, M., Merriman, J. A., Hutter, J., Lepenies, B., Rooijakkers, S. H. M., Malley, R.,
634 Walker, M. J., Shattil, S. J., Schlievert, P. M., Choudhury, B. and Nizet, V. (2014) The classical
635 lancefield antigen of group a Streptococcus is a virulence determinant with implications for
636 vaccine design. Cell Host Microbe. **15**, 729-740
637 45 Svetlana Zamakhaeva, C. T. C., Jeffrey S. Rush, Sowmya Ajay Castro, Alexander E.
638 Yarawsky, Andrew B. Herr, Nina M. van Sorge, Helge C. Dorfmueller, Gregory I. Frolenkov,
639 Konstantin V. Korotkov, Natalia Korotkova. (2020) Modification of cell wall polysaccharide
640 spatially controls cell division in Streptococcus mutans. bioRxiv
641 46 De, A., Liao, S., Bitoun, J. P., Roth, R., Beatty, W. L., Wu, H. and Wen, Z. T. (2017)
642 Deficiency of RgpG causes major defects in cell division and biofilm formation, and deficiency of

643 LytR-CpsA-Psr family proteins leads to accumulation of cell wall antigens in culture medium by
644 *Streptococcus mutans*. Appl Environ Microbiol. **83**
645 47 Rabinovich, M. L., Melnick, M. S. and Bolobova, A. V. (2002) The structure and
646 mechanism of action of cellulolytic enzymes. Biochemistry (Mosc). **67**, 850-871
647 48 McCabe, C., Zhao, X., Adney, W. S., Himmel, M. E. (2010) Energy Storage in Cellulase
648 Linker Peptides? Computational Modeling in Lignocellulosic Biofuel Production. **1052**, 119-134
649

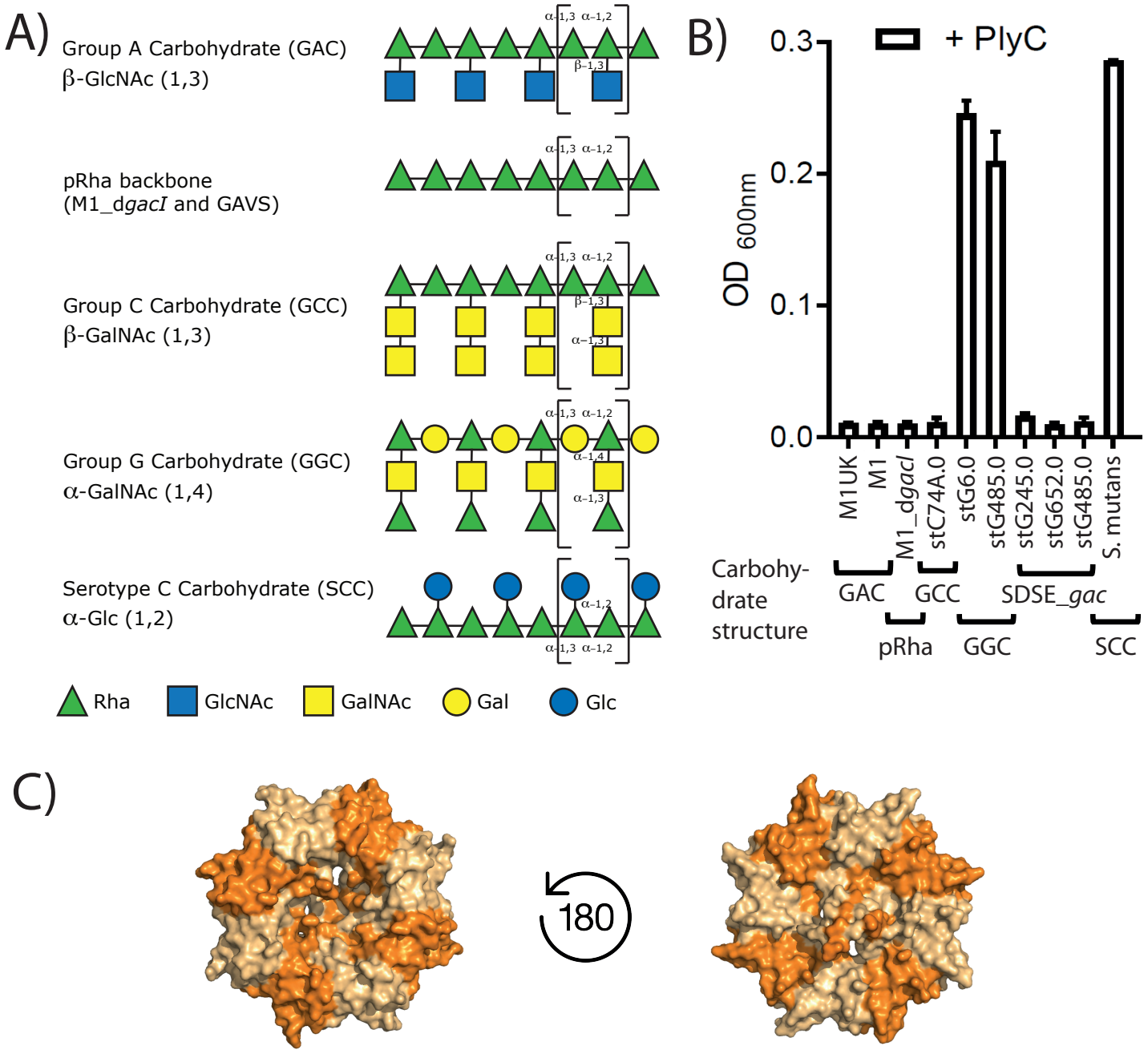

Fig 1)
A) Symbolic drawings of the carbohydrate structures of GAS (GAC), polyrhamnose backbone (pRha) in GAVS and M1_d*gacI*, GCS (GCC), GGS (GGC) and S. mutans (SCC). For simplicity, the reported glycerol phosphate of occasionally present on the GAC and SCC side chains have been omitted. Repeat units are marked with brackets. The pRha backbone with alternating (α1->2) and (α1->3) linkages, whilst the α1->2 rhamnose being decorated with a β1->3 side chain is a commonality among all PlyC susceptible strains. B) PlyC lysis of streptococcal pathogens. Group A and Group C Streptococci serotypes are susceptible to PlyC-mediated lysis. Group G Streptococci show limited susceptibility and S. mutans is resistant to PlyC lysis. Three SDSE isolates that produce the GAC instead of GGC (SDSE_*gac*) are susceptible to PlyC treatment. C) Bottom and top view of the WT PlyCB octameric ring structure with the GlyH and EAD domains omitted (4f78.pdb).

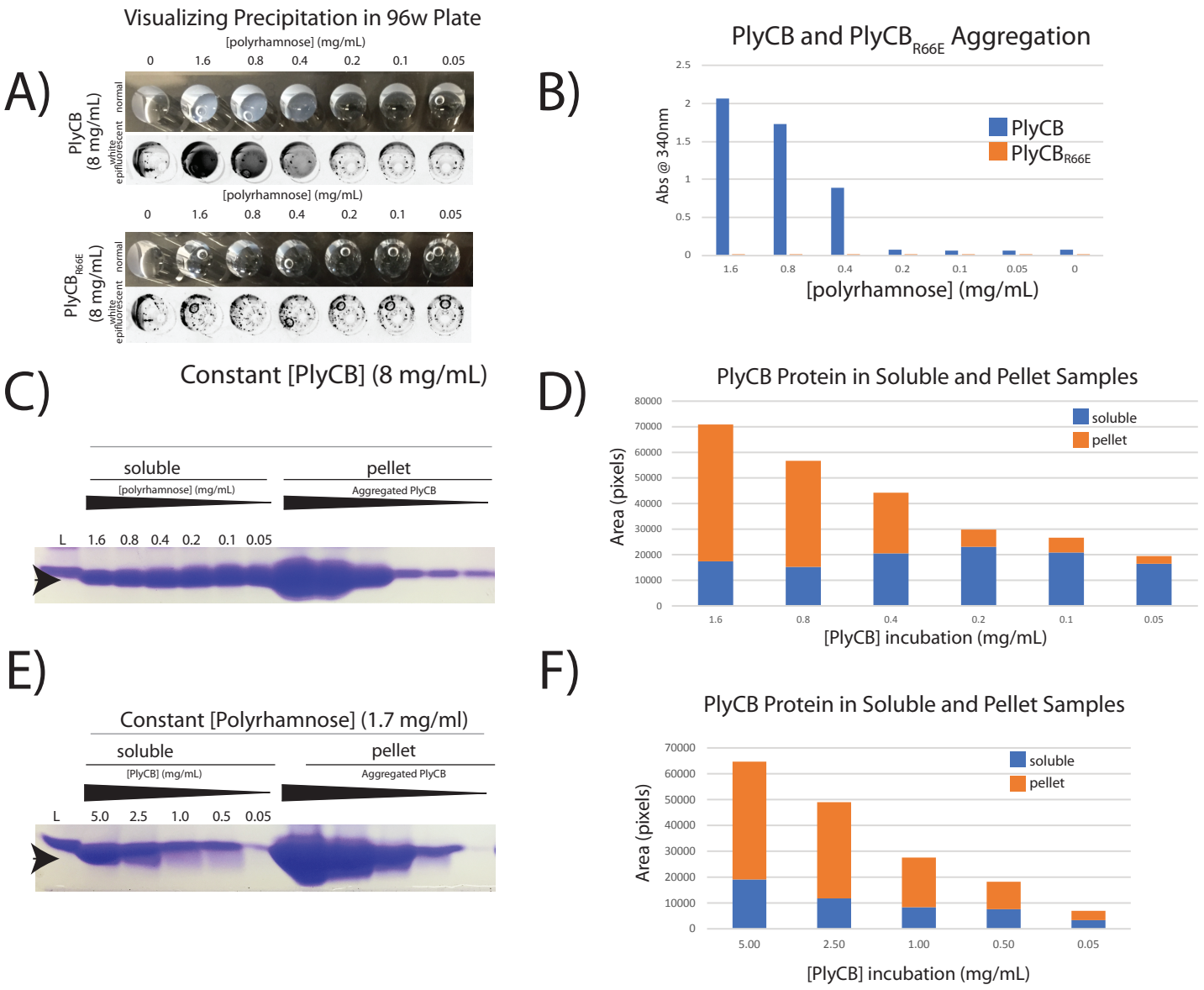

Fig 2)
Precipitation studies of purified PlyCB and GAC reveal direct interaction of PlyCB with GAC. A) The PlyCB concentration is kept constant whilst the GAC concentration is varied. Visible precipitate forms at the higher concentrations. B) The precipitate level is measured spectrophotometrically at 340 nm and compared to the mutant PlyCB_R66E, which does not bind the GAC. C) Coomassie stained and D) densitometry analysis of PlyCB protein from the supernatant fraction and aggregates (pellets). E, F) The same dose dependency is observed when the PlyCB concentration is varied. Arrowhead depicts PlyCB protein (8 kDa).

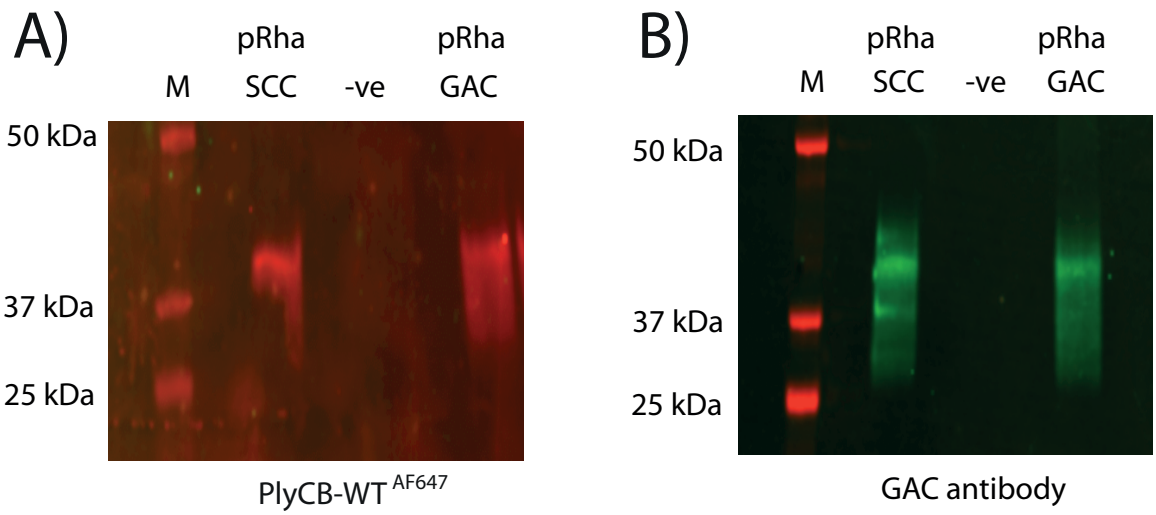

Fig 3)
Representation of immunoblot analysis of the cell lysate of E. coli expressing the SCC and GAC pRha and carrying an empty control plasmid (-ve). A) Blot was incubated with PlyCB-Alexa Fluor® 647 (PlyCB-WT-AF647). B) Probing the same samples with the GAC antibodies confirms the presence of GAC in the bands. Molecular mass markers are given in kDa.

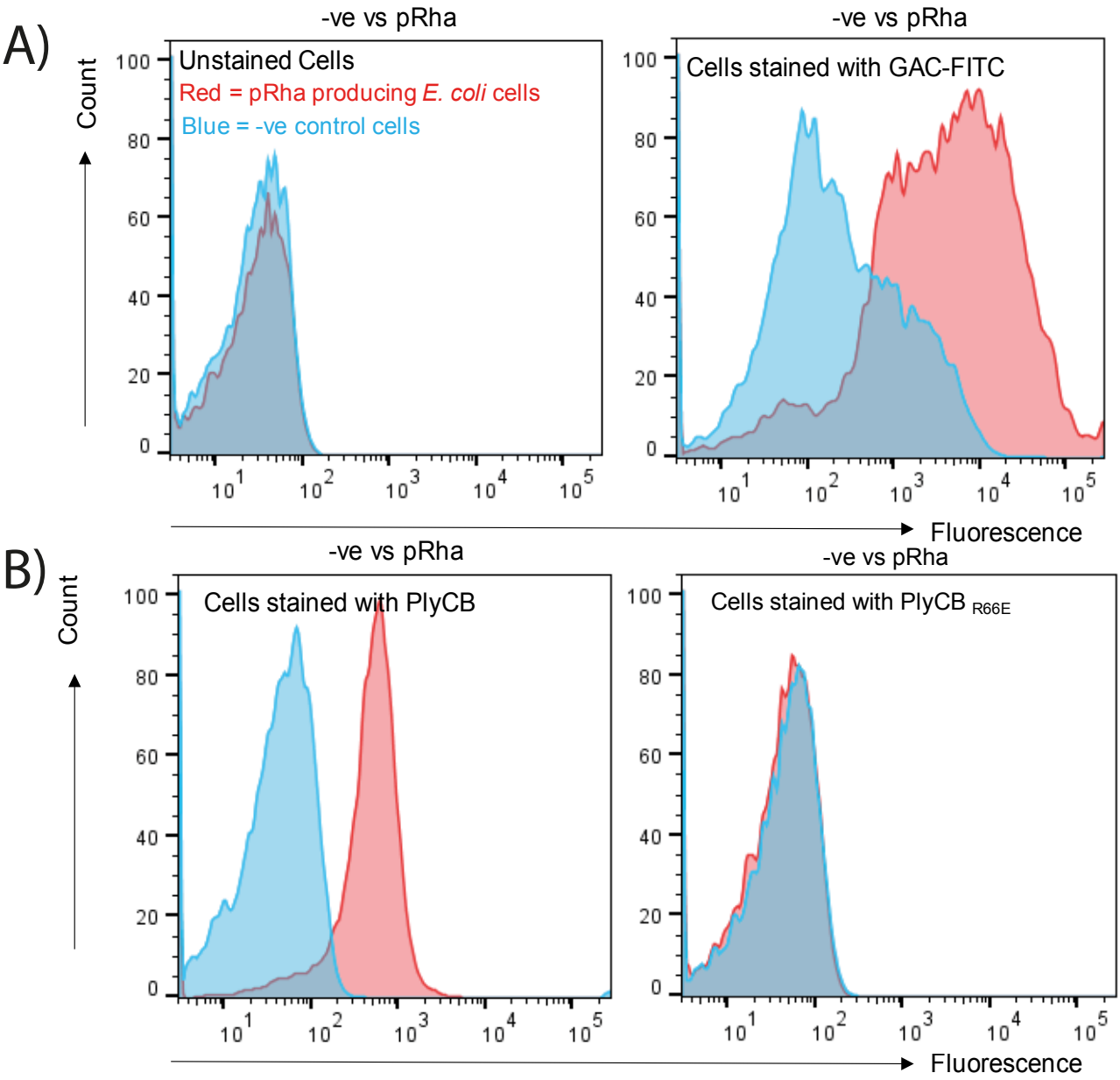

Fig 4)

PlyCB binding to *E. coli* cells were investigated by flow cytometry after labelling with $PlyCB_{WT}^{AF647}$ and $PlyCB_{R66E}^{AF645}$ mutant proteins. Blue: -ve control cells without pRha. Red: pRha producing *E. coli* cells. Representative histograms are shown. A) Left panel: unstained cells. Right panel: The anti-GAC antibodies (GAC-FITC) were used as a positive control to label the *E. coli* cells producing pRha. The antibodies do not bind to the E. coli cells carrying an empty plasmid (-ve). B) Left panel: $PlyCB_{WT}^{AF647}$ binds to the *E. coli* cells producing pRha, but not to the *E. coli* cells carrying an empty plasmid (-ve). Right panel: $PlyCB_{R66E}^{AF645}$ does not binds to the *E. coli* cells producing pRha.

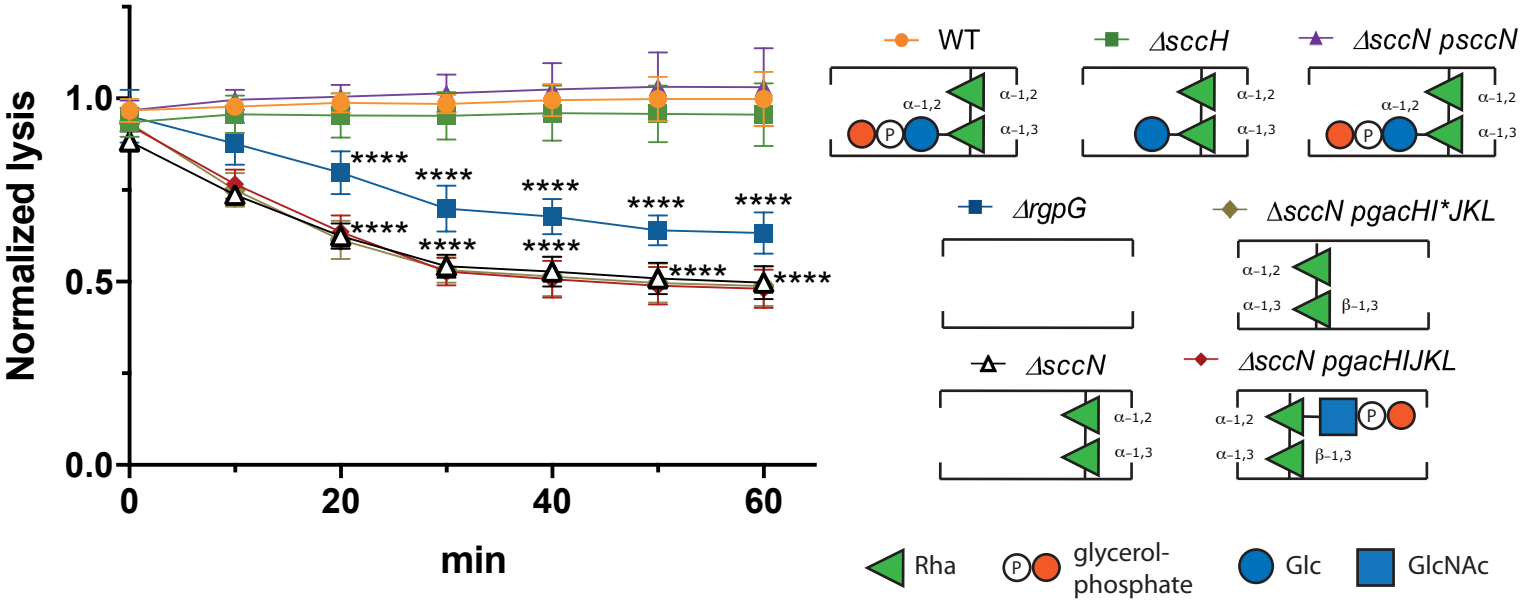

Fig 5)
The PlyC-mediated lysis of sacculi purified from *S. mutans* strains. The lysis was monitored after 10, 20, 30, 40, 50 and 60 min as a decrease in OD_{600}. Results are presented as a fold change in OD_{600} of the sacculi incubated with PlyC *vs.* the sacculi incubated without PlyC. Data points and error bars represent mean values of four biological replicates and standard deviation, respectively. P-values were determined by two-way ANOVA with Dunnett's multiple comparisons test.