

Biochemical Activities of *Streptococcus pneumoniae* Serotype 2 Capsular Glycosyltransferases and Significance of Suppressor Mutations Affecting the Initiating Glycosyltransferase Cps2E

David B. A. James,* Kanupriya Gupta, Jocelyn R. Hauser, Janet Yother

Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama, USA

The capsular polysaccharide (CPS) is essential for *Streptococcus pneumoniae* virulence. Its synthesis requires multiple enzymes, and defects that block completion of the pathway can be lethal in the absence of secondary suppressor mutations. In this study, we examined the functions of three capsular glycosyltransferases (Cps2F, Cps2G, and Cps2I) involved in serotype 2 CPS synthesis, whose deletions select for secondary mutations. We demonstrate that Cps2F is a rhamnosyltransferase that catalyzes addition of the third and fourth sugars in the capsule repeat unit, while Cps2G adds the fifth sugar (glucose). Addition of the terminal residue (glucuronic acid) could not be detected; however, activities of the other glycosyltransferases together with bioinformatic analyses suggest that this step is mediated by Cps2I. Most of the secondary suppressor mutations resulting from loss of these enzymes occur in *cps2E*, the gene encoding the initiating glycosyltransferase. Examination of the 69 *S. pneumoniae* serotypes containing Cps2E homologues yielded a consensus amino acid sequence for this protein and demonstrated that there is a highly significant association between the residues that are 100% conserved and those altered by suppressor mutations. Cps2E contains an extracytoplasmic loop whose function is unknown. Among our collection of mutants, six contained missense mutations affecting amino acids in the extracytoplasmic loop. These residues are highly conserved among *S. pneumoniae* Cps2E homologues, and mutations therein severely reduced CPS synthesis and Cps2E activity. The critical functions of these amino acids suggest a role for the Cps2E extracytoplasmic loop in initiation, and possibly regulation, of capsule synthesis.

Streptococcus pneumoniae (the pneumococcus) is an extracellular bacterial pathogen that produces a protective surface structure known as the capsular polysaccharide (CPS). The CPS is an essential component for virulence, influencing both nasopharyngeal colonization and systemic infection (1–8). CPS is also an immunodominant surface antigen that serves as the basis for *S. pneumoniae* serological typing and protective vaccines (3, 9–14). Over 90 CPS serotypes have been identified; all but 2 utilize the Wzy-dependent pathway to synthesize CPS, while the remaining serotypes (serotypes 3 and 37) utilize the synthase-dependent pathway (15–18). This study focused on the Wzy-dependent pathway, which was first identified as a mechanism of synthesis for lipopolysaccharide O antigen in Gram-negative bacteria (19–21) but is the principal mechanism for CPS synthesis in Gram-positive bacteria (18).

In the Wzy-dependent pathway, the capsule repeat unit is synthesized on the membrane acceptor undecaprenyl phosphate (Und-P) by multiple glycosyltransferases. Once completed, the repeat unit is transported across the cytoplasmic membrane by a Wzx translocase and polymerized by a Wzy polymerase (Fig. 1). As a model for studying the Wzy-dependent pathway in *S. pneumoniae*, our laboratory uses capsule serotype 2, which contains a hexasaccharide repeat unit comprised of glucose (Glc), rhamnose (Rha), and glucuronic acid (GlcUA) residues (22) that is polymerized to at least 1,000 to 2,000 repeat units, based on the molecular sizes of isolated polymers used in vaccine preparations (23) (Fig. 1). We previously identified and characterized the first two glycosyltransferases involved in synthesizing the serotype 2 repeat unit: the glucose-1-phosphate (Glc-1-P) transferase Cps2E, which initiates synthesis (24), and the β 1-4 rhamnosyltransferase Cps2T, which catalyzes the second and committed step (25).

There are a limited number of publications characterizing

pneumococcal glycosyltransferases involved in Wzy-dependent capsule synthesis. In *S. pneumoniae*, homologues of the initiating glycosyltransferase CpsE (also known as WchA) are present in at least 69 serotypes (15). The glycosyltransferase activities of these enzymes have been established, with enzyme functions demonstrated in serotypes 2, 8, 9v, and 14 (24, 26–28). CpsE is a member of the polyprenyl-phosphate hexose-1-phosphate transferase (PHPT) family (20), in which the *Salmonella enterica* galactose-1-phosphate transferase WbaP serves as the prototypic enzyme (29–31). Three distinct regions are present in PHPT enzymes: the N terminus, which in Cps2E contains 4 membrane-spanning domains; a central, extracytoplasmic loop; and a C-terminal cytoplasmic tail following a membrane-spanning domain (31–34). Sugar transfer is catalyzed by residues in the cytoplasmic tail, as independent studies have shown this region to be sufficient for glycosyltransferase activity (27, 31, 32, 35). The N-terminal region is suggested to be required for protein insertion or stability in the membrane (31). The function of the extracytoplasmic loop is less clear, but recent studies have suggested a role in modulation of chain length distribution (31).

Three other genes in the *cps2* locus—*cps2F*, *cps2G*, and *cps2I*—

Received 18 June 2013 Accepted 27 September 2013

Published ahead of print 4 October 2013

Address correspondence to Janet Yother, jyother@uab.edu.

* Present address: David B. A. James, Department of Microbiology, New York University School of Medicine, New York, New York, USA.

K.G. and J.R.H. contributed equally to this article.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JB.00715-13

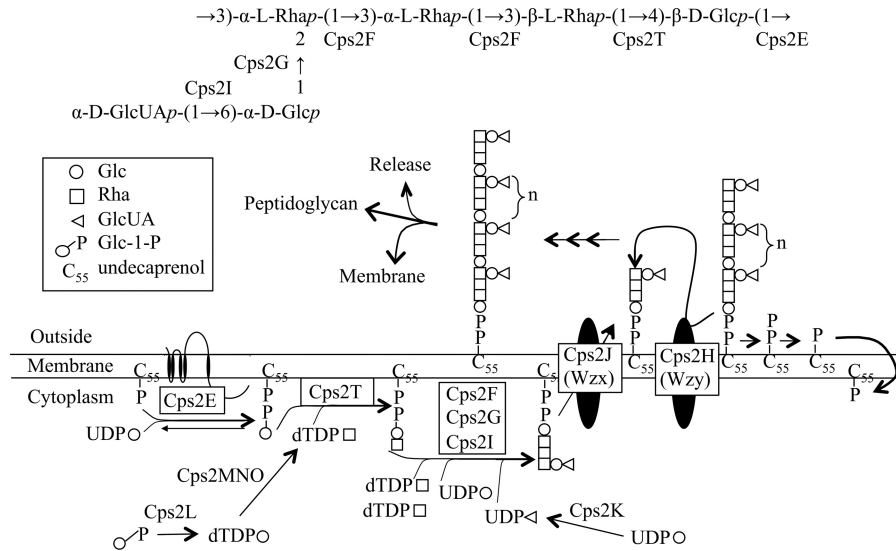


FIG 1 *S. pneumoniae* type 2 capsular polysaccharide repeat unit structure and Wzy-dependent pathway. (Top) Repeat unit structure. Letters below each glycosidic linkage represent the glycosyltransferase known or predicted to catalyze the linkage. (Bottom) Wzy-dependent model for synthesis of the type 2 repeat unit. The number of repeat units (n) is estimated to be 1,000 to 2,000, based on the molecular size of isolated polymer used in vaccine preparations (23). Figure adapted from references 25 and 55.

are predicted to encode the glycosyltransferases required for completion of the repeat unit (Fig. 1) (25, 36, 37), but biochemical evidence supporting their functions is lacking. In previous reports, we demonstrated that deletion of any one of these genes, or of the UDP-glucose dehydrogenase encoded by *cps2K*, the Wzx translocase (*cps2J*), or the Wzy polymerase (*cps2H*), was lethal in the absence of secondary mutations, the majority of which mapped to *cps2E* (25, 34). Each of the primary mutations affected a point in the pathway past the committed step (Cps2T-catalyzed addition of the first Rha). The lethality was attributed to either destabilization of the membrane by the buildup of incomplete repeat units or sequestration of Und-P in the capsule pathway and away from essential pathways such as peptidoglycan synthesis. Other reports have described similar suppressor mutations, some of which occurred in the initiating glycosyltransferase, when primary mutations blocked the completion of polysaccharide pathways in *Pseudomonas aeruginosa* lipopolysaccharide (LPS) Wzx flippase mutants (38), *Xanthomonas campestris* xanthan gum mutants (39), and *Salmonella enterica* serovar Typhimurium LPS mutants that failed to polymerize O-antigen subunits (30). To date, we have obtained over 30 unique suppressor mutations in *cps2E* that are distributed throughout the C-terminal cytoplasmic tail and the central extracytoplasmic loop (25, 34). Suppressor mutations mapping to the C-terminal region can be interpreted as a direct inhibition of Glc-1-P transferase activity. Suppressor mutations that map to the extracytoplasmic loop, however, are more difficult to interpret but must have a function in limiting the accumulation of incomplete repeat units.

In this report, we provide biochemical evidence for the previously uncharacterized glycosyltransferases Cps2F and Cps2G that, together with bioinformatic analyses, completes the transferase specificities for the type 2 CPS repeat unit of *S. pneumoniae*. We also further characterize our novel collection of secondary mutations in *cps2E* and demonstrate that residues in the extracytoplasmic loop are critical for repeat unit initiation.

MATERIALS AND METHODS

Bacterial strains and growth conditions. *S. pneumoniae* strains (Table 1) were grown at 37°C in THY (Todd-Hewitt broth supplemented with 0.5% yeast extract) or on blood agar plates (BBL [Difco] or blood agar base [Becton and Dickinson] containing 3% defibrinated sheep blood [Colorado Serum Company]). Broth cultures were grown statically at 37°C. Plate cultures were incubated at 37°C in a candle jar. *Escherichia coli* strains (Table 1) were grown at 37°C in L broth (1% tryptone, 0.5% yeast extract, 0.5% NaCl, and 0.1% glucose) with shaking or on L agar plates (15 g of agar/liter of L broth). Media were supplemented with the following antibiotics as appropriate: ampicillin (Ap; 100 µg/ml), erythromycin (Em; 15 µg/ml for *E. coli* DB11 and 0.3 µg/ml for *S. pneumoniae*), and kanamycin (Km; 250 µg/ml).

Cloning of glycosyltransferase mutant enzymes. Sequences for *cps2E* with the mutations *cps2E*^{5718A→G} (H258R), *cps2E*^{5529G→T} (A162S), *cps2E*^{5544G→T} (D167Y), *cps2E*^{5539C→T} (L200F), and *cps2E*^{5632T→G} (V196G) were PCR amplified from the chromosomal DNA of *S. pneumoniae* strains DJ916, DJ921, BX547, BX552, and BX605, respectively. The primer pair used in PCRs was E8 (CATATGAATGGAAAAACAGTA AAGTC) and E11 (CTCGAGCTACTTCGCTCCATCTCTC), which amplifies a 1.2-kb fragment containing the full-length *cps2E* (24). PCR products were cloned into pET-20b and transformed into chemically competent *E. coli* BL21-AI. The resulting strains DJ252 (Cps2E^{H258R}), DJ253 (Cps2E^{A162S}), DJ254 (Cps2E^{D167Y}), DJ269 (Cps2E^{L200F}), and DJ270 (Cps2E^{V196G}) were used for expression of proteins for glycosyltransferase assays, as described below.

Introduction of *cps2E* point mutations into *S. pneumoniae* D39. To facilitate identification of isolates transformed with *cps2E* point mutations, allelic exchange was used to insert a kanamycin resistance-encoding (Km^r) cassette (*aphA-3*) into the native *cps2E*. The *aphA-3* gene was PCR amplified from the pneumococcal shuttle vector pSF151 (40) using primer pair DJ-03/DJ-04 and cloned into pGEM-T Easy cloning vector, resulting in pDJ271 (all plasmids were isolated and maintained in *E. coli* DH5α or DB11). Following digestion of pDJ271 with BstBI, the released *aphA-3*-containing fragment was ligated into BstBI-digested pBX146. The latter plasmid contains the full-length *cps2E* and partial upstream and downstream sequences (*cps2D* and *cps2T*, respectively) cloned into the pneumococcal suicide vector pJY4164, which encodes Em resistance

TABLE 1 Strains and plasmids used in this study

Strain or plasmid	Properties ^a	Reference or source
<i>S. pneumoniae</i>		
BX547	D39 $\Delta cps2K$ Cps ⁺ ; cps2E ^{5544 G→T} (D167Y) Km ^r	34
BX552	D39 $\Delta cps2H$ Cps ⁻ cps2E ^{5539 C→T} (L200F); Km ^r (in the original publication, this mutant was incorrectly identified as L199F)	34
BX605	D39 $\Delta cps2K$ Cps ⁺ cps2E ^{5632T→G} (V196G) Km ^r	34
BX612	D39 $\Delta cps2K$ Cps ⁺ cps2E ^{5539T→G} (V165G) Km ^r	34
D39	Type 2 parent strain; Cps ⁺	57
DJ916	D39 $\Delta cps2F$ Cps ⁻ cps2E ^{5718 A→G} (H258R) Km ^r	25
DJ918	D39 $\Delta cps2G$ Cps ⁺ cps2E ^{6059G→C} (M338I) Km ^r	25
DJ919	D39 $\Delta cps2G$ Cps ⁺ cps2E ^{6964C→A} (R307S) Km ^r	25
DJ921	D39 $\Delta cps2G$ Cps ⁺ cps2E ^{5529G→T} (A162S) Km ^r	25
JK901	pJK101 X D39; Cps ⁻ ; <i>aphA-3</i> inserted into <i>cps2E</i> ; Km ^r	This study
JK903	pBX146 X JK901; <i>cps2E</i> repair; Cps ⁺	This study
JK906	pDJ238 X JK901; <i>cps2E</i> ^{5632T→G} (V196G) Cps ^r	This study
JK907	pDJ241 X JK901; <i>cps2E</i> ^{5718A→G} (H258R) Cps ^r	This study
KA1522	D39 $\Delta cps2E$ Cps ⁻	24
WU2	Type 3 parent strain; Cps ⁺	58
<i>E. coli</i>		
BL21-AI	F ⁻ <i>ompT hsdSB</i> (r _B ⁻ m _B ⁻) <i>gal dcm araB::T7 RNAP-tetA</i>	Invitrogen
DB11	<i>met thi gal hasdR nal rif</i> ; plasmid-cured derivative of V854 from reference 59	59
DH5 α	F ⁻ Φ 80lacZ Δ M15 (Δ (lacZYA-argF) U169 <i>recA1 endA1 hsdR17</i> (r _K ⁻ m _K ⁻) <i>phoA supE44</i> λ ⁻ <i>thi-1 gyrA96 relA1</i>	Invitrogen
DJ005	BL21-AI expressing Cps2I; Ap ^r	25
DJ009	BL21-AI expressing Cps2T; Ap ^r	25
DJ011	BL21-AI expressing Cps2G; Ap ^r	25
DJ204	BL21-AI expressing Cps2F, replaced GTG start with ATG; Ap ^r	25
DJ238	DB11 (pDJ248); used to transform <i>cps2E</i> ^{5718A→G} (H258R)	
DJ241	DB11 (pDJ241); used to transform <i>cps2E</i> ^{5632T→G} (V196G)	
DJ252	BL21-AI (pDJ252) expressing Cps2E ^{H258R} ; Ap ^r	This study
DJ253	BL21-AI (pDJ253) expressing Cps2E ^{A162S} ; Ap ^r	This study
DJ254	BL21-AI (pDJ254) expressing Cps2E ^{D167Y} ; Ap ^r	This study
DJ269	BL21-AI (pDJ269) expressing Cps2E ^{L200F} ; Ap ^r	This study
DJ270	BL21-AI (pDJ269) expressing Cps2E ^{V196G} ; Ap ^r	This study
DJ271	DH5 α (pDJ271); <i>aphA</i> -containing clone; Km ^r	This study
JK101	DB11 (pJK101); contains <i>aphA-3</i> -disrupted <i>cps2E</i> ; Km ^r	This study
KJ4152	BL21-AI (pKJ4152) expressing full-length Cps2E; Ap ^r	25
RC124	BL21-AI (pET20b) vector control; Ap ^r	24
Plasmids		
pBX146	pJY4164/ <i>cps2E</i> ; amplified from D39 using primer pair cps2-D3/cps2-T1; contains full-length parental <i>cps2E</i> . This plasmid is the same as the published pBX145.	34
pET20b	Protein expression vector; Ap ^r	Novagen
pDJ238	pBX146/ <i>cps2E</i> ^{5718A→G} (H258R) amplified from DJ916 using primers E8/E11 ^b	This study
pDJ241	pBX146/ <i>cps2E</i> ^{5632T→G} (V196G) amplified from BX605 using primers E8/E11 ^b	This study
pDJ252	pET-20b/ <i>cps2E</i> ^{5718 A→G} (H258R) amplified from DJ916 using primers E8/E11 ^b	This study
pDJ253	pET-20b/ <i>cps2E</i> ^{5529G→T} (A162S) amplified from DJ921 using primers E8/E11 ^b	This study
pDJ254	pET-20b/ <i>cps2E</i> ^{5544 G→T} (D167Y) amplified from BX547 using primers E8/E11 ^b	This study
pDJ269	pET-20b/ <i>cps2E</i> ^{5539 C→T} (L200F) amplified from BX552 using primers E8/E11 ^b	This study
pDJ270	pET-20b/ <i>cps2E</i> ^{5632 T→G} (V196G) amplified from BX605 using primers E8/E11 ^b	This study
pDJ271	pGEM-T Easy/ <i>aphA-3</i> -containing fragment amplified using primer pair DJ-03/DJ-04	This study
pJK101	pBX146/ <i>aphA-3</i> from BstBI-digested product of pDJ271	This study
pJY4164	<i>S. pneumoniae</i> suicide vector; Em ^r	41
pKJ4152	pET-20b/ <i>cps2E</i> amplified from <i>S. pneumoniae</i> D39 using primers E8/E11 ^b	25

^a *cps2E* superscripts indicate locations of mutations based on GenBank accession no. AF026471. Amino acid changes are indicated in parentheses. Cps^r, reduced capsule levels; →, nucleotide change; Ap^r, ampicillin resistant.

^b Primers are described in Materials and Methods.

(Em^r) (41). BstBI cuts pBX146 at a single site within *cps2E*, allowing introduction of the *aphA-3* fragment. The resulting construct, pJK101, was used to transform competent D39 (42). Km^r isolates were demonstrated to be CPS negative and Em sensitive (Em^s). The latter property indicates that Km^r was the result of allelic exchange and not single crossover resulting in integration of the entire clone into the recipient chromosome. The Km^r isolate JK901 was used in subsequent studies.

Constructs containing the *cps2E* suppressor mutations encoding V196G and H258R were generated by PCR amplification of chromosomal DNA from strains DJ916 and BX605, respectively, using primer pair E8/E11 (primer sequences listed above) to amplify the full-length *cps2E*. The fragments were cloned into the pGEM-T Easy cloning vector, and the subsequent clone was then digested with MscI/BlpI. The released fragments were ligated to MscI/BlpI-digested pBX146. Digestion with MscI/BlpI releases a 751-bp internal fragment of *cps2E* that encompasses the sites of the mutations. The mutant plasmids resulting from these ligations (pDJ238, V196A, and pDJ241, H258R), along with the parent pBX146, were transformed into competent JK901 (*cps2E::aph-3* Km^r), and reaction mixtures were plated on blood agar in the absence of antibiotic selection. Individual colonies were screened for Km sensitivity by patching on blood agar with and without Km. Sensitive isolates were tested for Em sensitivity to confirm that the entire plasmid had not integrated into the chromosome. The resulting strains were confirmed to have only the expected mutations (JK906, V196, and JK907, H258) or the parental sequence (JK903) by DNA sequencing of *cps2E*. The frequencies of obtaining these mutants were 0.3% (13/4,000) for JK906, 0.1% (6/6,000) for JK907, and 0.7% (27/4,000) for JK903. Sequencing was performed by the UAB Genomics Core Facility of the Hefflin Center for Genomic Science.

Glycosyltransferase expression and membrane isolation. The previously constructed *E. coli* strains expressing Cps2E (KJ4152), Cps2T (DJ009), Cps2F (DJ204), Cps2G (DJ011), and Cps2I (DJ005) were cultured and processed for membrane isolation as previously described (25). None of these constructs or those for other glycosyltransferases contained tags. For expression, 250-ml cultures were grown with the appropriate antibiotic selection to mid-exponential phase (cell density of approximately 4.5×10^8 CFU/ml) and then induced with 0.2% arabinose for 30 min (Cps2E and derivatives) or 4 h (all others). Cultures were centrifuged, and the cells were washed and lysed as previously described (25), followed by centrifugation at $100,000 \times g$ to collect membranes. The final membrane pellet was suspended in 500 μ l Tris-acetate (100 mM, pH 7.5) containing 10% glycerol and 1 mM dithiothreitol (DTT; buffer A) and stored at -80°C . Frozen membranes were thawed and vigorously vortexed, and the homogenized solution was used to determine total membrane protein concentrations using the Bradford Bio-Rad protein dye assay. Replicate samples (at least duplicate) were assayed. Membrane samples from the parent and mutant strains demonstrated similar levels of a non-capsule-related protein, indicating homogeneity of the solutions (see Results).

S. pneumoniae membranes were isolated based on previously described procedures (43, 44). Fifty-milliliter cultures were grown at 37°C in THY to a density of 3×10^8 CFU/ml. Cultures were centrifuged at $10,000 \times g$ for 20 min. The cells were washed once in 50 ml of phosphate-buffered saline (137 mM NaCl, 2.7 mM KCl, 5.4 mM Na₂HPO₄·7H₂O, 1.8 mM KH₂PO₄ [pH 7.4]) and then suspended in 500 μ l of protoplast buffer (20% sucrose, 50 mM Mg₂SO₄, 50 mM Tris-HCl [pH 7.4]). Samples were incubated overnight at room temperature with gentle rotation (the *S. pneumoniae* autolysin LytA is active under these conditions [44]). Protoplast formation was confirmed using phase-contrast light microscopy. The protoplasts were centrifuged at $16,000 \times g$ for 10 min and washed once in 500 μ l of protoplast buffer. Protoplasts were osmotically lysed by suspension in 1 ml of sterile water containing 10 mM EDTA, 1 mM DTT, and 1 mM phenylmethanesulfonyl fluoride, with incubation at room temperature for 30 min. Lysis was confirmed by light microscopy. Lysed cells were centrifuged at 16,000 or $100,000 \times g$ (with equivalent results) for 20 min. The membrane-containing pellets were washed once in 1 ml of buf-

fer A and suspended in 100 μ l of buffer A. Total membrane protein concentrations were quantified using the Bradford Bio-Rad protein dye assay, as described above.

Synthesis, purification, and quantification of dTDP-Rha. dTDP-Rha was synthesized in two steps essentially as previously described (25), with minor modifications to enzyme concentration and incubation time. Briefly, dTDP-Glc was synthesized in a 350- μ l reaction mixture containing 50 mM HEPES (pH 7.6), 30 mM MgCl₂, 3.4 mM dTTP, 3.4 mM Glc-1-P, 2 U of inorganic pyrophosphatase, and 2 μ M Cps2L. After incubation for 90 min at 37°C, the proteins were removed by filtration using Amicon Ultra 10-kDa centrifugal filters (Millipore). The filtrate was supplemented with 0.1 mM NAD⁺, 6 mM NADH, 1 μ M Cps2N, 1 μ M Cps2M, and 1 μ M Cps2O in a final volume of 350 μ l. The reaction mixture was incubated for 30 min at 37°C, the proteins were removed by filtration, and dTDP-Rha was separated by high-performance liquid chromatography (HPLC) as previously described (25).

Glycosyltransferase assays. Glycosyltransferase assays to assess the functions of Cps2F, Cps2G, and Cps2I utilized recombinant enzymes expressed separately in *E. coli*. For each enzyme, membranes containing 3 μ g of total protein were incubated for 1 h at 10°C in reaction mixtures containing 10 mM MnCl₂, 1 mM DTT, 0.008% NP-40 (in experiments utilizing multiple glycosyltransferases), 0.075 μ Ci of UDP-[¹⁴C]Glc (300 mCi/mmol), and, where indicated, 0.1 mM dTDP-Rha and/or 0.1 mM UDP-GlcUA. The final volume was brought to 75 μ l with Tris-acetate (100 mM, pH 7.6). Reactions were stopped with 1 ml of 2:1 chloroform-methanol, and the organic phase was extracted three times with pure solvent upper phase (PSUP; 1.5 ml of chloroform, 25 ml of methanol, 23.5 ml of H₂O, and 0.183 g of KCl). Glycolipids were either preserved for later use as described for Und-P-P-[¹⁴C]-Glc-Rha-Rha-Rha (see below) or subjected to mild-acid hydrolysis by suspending the dried organic phase in 80 μ l of 20 mM HCl and incubating it at 70°C for 20 min. Hydrolyzed samples were dried, suspended in 10 μ l of chloroform-methanol (1:1), and applied as spots on silica-coated thin-layer chromatography (TLC; Whatman) plates. TLC plates were chromatographed for 8 h in butanol-ethanol-water (5:3:2). Dried plates were exposed to a phosphor screen (General Electric Healthcare) for 15 h, and bands were visualized in a Storm 820 phosphorimager (General Electric Healthcare).

Assays to quantify Cps2E activities of the parent and mutant proteins were performed as previously described (24). Except as noted, *S. pneumoniae* or *E. coli* membranes containing the indicated amounts of total protein (see figure legends) were incubated for 1 h at 10°C in 75- μ l reaction volumes containing 10 mM MnCl₂, 1 mM DTT, and 0.075 μ Ci of UDP-[³H]Glc (43.8 Ci/mmol) in Tris-acetate (100 mM, pH 7.6). The reaction was stopped, and glycolipids were extracted and dried as described above for assays involving *E. coli* membranes. Incorporated [³H]Glc was analyzed by liquid scintillation counting.

Synthesis of Und-P-P-[¹⁴C]Glc-Rha-Rha-Rha (Cps2F product) and assay of Cps2G activity in the absence of other glycosyltransferases. Reaction mixtures of membranes containing Cps2E, Cps2T, and Cps2F were incubated as described above to synthesize the Cps2F product. Extracted glycolipids were pooled, dried, and suspended in 100 μ l of Tris-acetate (100 mM; pH 7.6) containing 0.1% NP-40. This solution was used as the source of Und-P-P-[¹⁴C]Glc-Rha-Rha-Rha in subsequent glycosyltransferase reactions. Membranes (3 μ g of total protein) from DJ011 containing Cps2G or from RC124, a vector-only *E. coli* strain, were incubated in reaction mixtures containing 0.008% NP-40, 10 mM MnCl₂, 1 mM DTT, 25 μ l of the Und-P-P-[¹⁴C]Glc-Rha-Rha-Rha solution, and 0.1 mM UDP-Glc. Reaction products were visualized as described above.

Bioinformatic analysis of the Cps2E homologues in *S. pneumoniae* and other bacteria. Amino acid sequences from 94 serotypes (90 listed in reference 15 and the remaining 4 in references 45 and 46) were analyzed for Cps2E homologues, which occur in 69 serotypes. These amino acid sequences were aligned by ClustalW (Geneious, version 5.6.4), and a consensus amino acid sequence was generated. Using Fisher's exact test, a significant correlation ($P < 0.01$) was observed between amino acids that are 100% conserved

among *S. pneumoniae* Cps2E homologues and those selected as suppressor mutations. A protein basic local alignment search tool (BLAST [<http://blast.ncbi.nlm.nih.gov/>]) was used to identify Cps2E homologues in other bacteria. Homologous proteins from *Clostridium phytofermentans* ISDg (GenBank accession number YP_001558318), *Clostridium bolteae* (WP_002576468), *Streptococcus oralis* (BAD22619), *Streptococcus mitis* (WP_001013652), and *Streptococcus parasanguinis* (WP_003009077) were among those identified. Amino acid sequences from these homologues, together with WbaP from *Salmonella enterica* (YP_008266333) were aligned by ClustalW as described above.

RESULTS AND DISCUSSION

Functional analyses of glycosyltransferases encoded in the *cps2* locus. We previously reported that Cps2E and Cps2T catalyze the first and second steps, respectively, in type 2 CPS synthesis (24, 25). In addition to these enzymes, the *cps2* locus encodes three additional predicted glycosyltransferases (Cps2F, Cps2G, and Cps2I) that are expected to catalyze the remaining four glycosidic linkages that complete the capsule repeat unit. The observation that three enzymes might catalyze four linkages suggested that a single glycosyltransferase encoded within the *cps* locus is capable of adding two sugar residues. We therefore sought to demonstrate the enzyme activities of Cps2F, Cps2G, and Cps2I and also clarify the functions of the gene products whose deletions select for secondary mutations in *cps2E* (25).

In *in vitro* glycosyltransferase assays, we utilized individual membrane preparations obtained from *E. coli* expressing *S. pneumoniae* glycosyltransferase Cps2F, Cps2G, or Cps2I. The absence of functionally equivalent enzymes in *E. coli* was demonstrated by the lack of activity when using *E. coli* membranes from a strain containing only the plasmid vector, as described in the sections below. As previously reported (25), each of the *S. pneumoniae* glycosyltransferases is membrane associated when expressed in *E. coli* despite lacking membrane-spanning domains. As for Cps2T (25), *in vitro* enzyme activity required the addition of nonionic detergent (0.008% NP-40) at a concentration well below the critical micelle concentration. This requirement may reflect a need for temporal disassociation of enzymes from the membrane and access to membrane-bound glycolipid acceptors, possibly as part of the same micelle.

(i) Cps2F. To identify the enzyme that catalyzes the third step in CPS synthesis, i.e., the second rhamnosyltransferase, we combined Cps2E-containing membranes in reactions with membranes containing Cps2T, nucleotide sugar substrates, and membranes containing the remaining glycosyltransferases. Glycolipids formed in these reactions were extracted and the saccharide moiety was liberated by mild-acid hydrolysis and then separated by TLC. In the presence of UDP-[¹⁴C]Glc and dTDP-Rha, only the reaction mixture that contained Cps2F resulted in the formation of a new [¹⁴C]Glc-labeled product (Fig. 2), indicating that Cps2F catalyzes the third step in CPS synthesis. The Cps2F-dependent product was consistently generated in several independent reactions (data not shown).

The serotype 2 capsule has two identical inverting linkages (Rha- α -1,3-Rha- α -1,3) (22) (Fig. 1); however, Aanensen et al. identified only one gene (*cps2F*, also annotated as *wchG*) in the *cps* locus that was predicted to encode an inverting transferase (36). This observation suggests that Cps2F may add both Rha residues. Using our standard TLC separation system, we were unable to distinguish two new products (Fig. 2). Alterations in the method, including adjustments to the polarity of the mobile phase, length

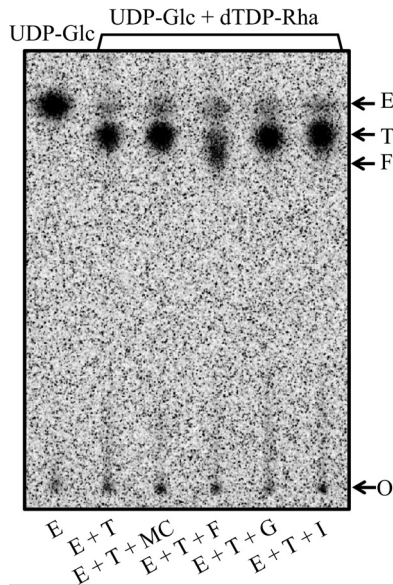


FIG 2 Cps2F rhamnosyltransferase activity. Cps2E-containing membranes (3 μ g of total protein) were incubated with membranes containing the indicated glycosyltransferases (3 μ g of total protein) or membrane controls (MC; 3 μ g of total protein from vector-only *E. coli* strain RC124) and UDP- 14 C]Glc only or UDP- 14 C]Glc and dTDP-Rha. Reaction mixtures (75- μ l total volume) were incubated for 1 h at 10°C and contained 0.008% NP-40, 10 mM MnCl₂, 1 mM DTT, 0.075 μ Ci of UDP- 14 C]Glc (300 mCi/mmol), and, where indicated, 0.1 mM dTDP-Rha. Glycolipids were extracted, hydrolyzed, and separated by TLC as described in Materials and Methods. Lanes show enzymes in the reaction mixtures: E, Cps2E; T, Cps2T; MC, membrane control; F, Cps2F; G, Cps2G; and I, Cps2I. Arrows indicate the locations of the origin (O) and the products formed by Cps2E, Cps2T, and Cps2F. The Cps2E product migrates equivalently to a 14 C]glucose standard.

of run, silica matrix composition, and separation in 2 dimensions, did not greatly increase the separation of the Cps2F and Cps2T products or allow for resolution of a second Cps2F-dependent product (data not shown). Rapid addition of two sugars by Cps2F may not permit distinction of two separate products, and subsequent data demonstrating Cps2G activity (described below) provided evidence for a Cps2F-dependent addition of two consecutive α 1-3 Rha residues.

(ii) **Cps2G.** To determine which enzyme catalyzes the next step in CPS synthesis, Cps2E-, Cps2T-, and Cps2F-containing membranes were combined with membranes containing Cps2G or Cps2I and UDP- 14 C]Glc and dTDP-Rha. In this case, only reaction mixtures with Cps2G resulted in a new 14 C]Glc-labeled product (Fig. 3A), indicating that Cps2G activity follows Cps2F activity. Since both UDP- 14 C]Glc and dTDP-Rha were added to the reaction mixture, we next sought to identify the sugar specificity of Cps2G. Here, reaction mixtures with membranes containing Cps2E, Cps2T, and Cps2F were used to synthesize the Cps2F lipid-linked reaction product, which was then extracted and incubated with UDP-Glc or dTDP-Rha, and membranes with or without Cps2G. As shown in Fig. 3B, Cps2G was capable of adding Glc, but not Rha, to the Cps2F product, indicating that Cps2G is the α 1-2 glucosyltransferase. Although we have not determined the identities of the products formed in these reactions through other biochemical or structural means, these results provide strong evidence for the identity of the Cps2F reaction product as Und-P-P-Glc-Rha-Rha-Rha and indicate that Cps2F sequen-

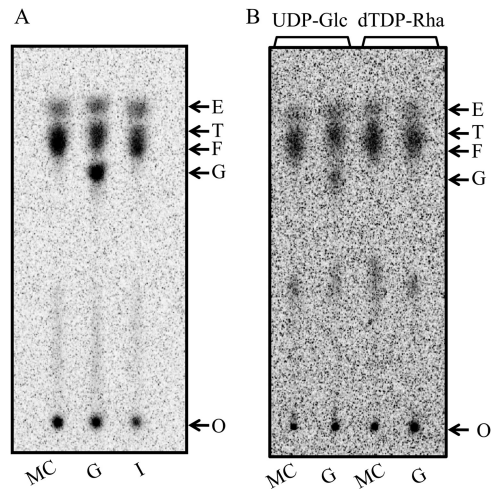


FIG 3 Cps2G glucosyltransferase activity. (A) Cps2E-, Cps2T-, and Cps2F-containing membranes (3 μ g of total protein each) were incubated with UDP- 14 C]Glc, dTDP-Rha, and membranes containing Cps2G, Cps2I, or a membrane control (MC; 3 μ g of total protein from vector-only *E. coli* strain RC124). Reaction conditions are as indicated for Fig. 2. (B) Cps2G activity in the absence of Cps2E, Cps2T, and Cps2F. The extracted Cps2F glycolipid product (Und-P-P- 14 C]Glc-Rha-Rha-Rha) was incubated with membranes containing Cps2G or membrane controls in the presence of UDP-Glc or dTDP-Rha. Reaction mixtures contained 3 μ g of total membrane protein, 0.008% NP-40, 10 mM MnCl₂, 1 mM DTT, 25 μ l of the Und-P-P- 14 C]Glc-Rha-Rha-Rha solution (see Materials and Methods), and 0.1 mM dTDP-Rha or 0.1 mM UDP-Glc. Reactions were processed as described above. Lanes show enzymes in the reaction mixtures: MC, membrane control; G, Cps2G; and I, Cps2I. Arrows indicate the locations of the origin (O) and the products formed by Cps2E, Cps2T, Cps2F, and Cps2G.

tially adds two Rha- α 1,3 residues, followed by the addition of Glc by Cps2G.

A mechanism for sequential addition of multiple sugars by a bacterial glycosyltransferase has been reported for PglH from *Campylobacter jejuni*, which adds three α 1-4 *N*-acetylgalactosamine (GalNAc) residues (47). For PglH, an initial processive synthesis is followed by a dissociative mechanism as products accumulate. Limitation to only three sugar additions is proposed to result from increased binding affinity of products with increasing size, as well as an inability of larger products to move through the active site of the enzyme. PglH contains five copies of the nine-amino-acid motif EX₇E, but only one of these was critical for sequential addition of GalNAc residues, suggesting the presence of a single active site (47). The EX₇E motif is common among glycosyltransferases (48). It may be essential for activity, as in PglH, or important but not essential, as in PelF from *Pseudomonas aeruginosa* (49). Similarly, the presence of multiple EX₇E motifs could be indicative of a multidomain enzyme with more than one active site, as for enzymes in the *E. coli* O8 and O9a polymannose O-polysaccharide biosynthetic pathways (50). The EX₇E motif occurs once in Cps2F (E₁₀₀/E₁₀₈) and several times in other glycosyltransferases of the type 2 pathway (Cps2E, 5 sites; Cps2T, 3 sites; Cps2G, 2 sites; and Cps2I, 2 sites). If this motif is essential for Cps2F activity, the presence of only one EX₇E would suggest a single active site for the enzyme. Determination of the importance of the EX₇E motif, the number of active sites, and whether control of the number of Rha residues added occurs by a PglH-like counting mechanism awaits further mutational and biochemical studies.

(iii) **Cps2I.** To test the ability of Cps2I to add GlcUA, the final

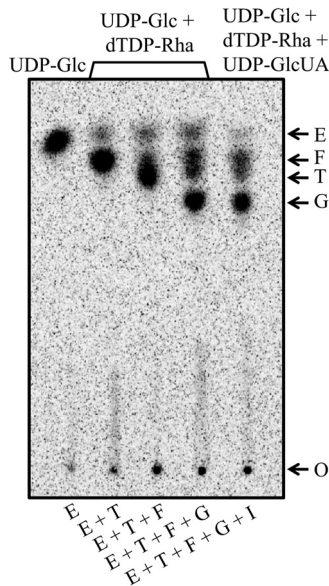


FIG 4 Assay for Cps2I glycosyltransferase activity. Cps2E-containing membranes (3 μ g of total protein) were incubated with membranes containing the indicated glycosyltransferases (3 μ g of total protein) and nucleotide sugars. Reactions were processed as described for Fig. 2 and contained 0.008% NP-40, 10 mM MnCl₂, 1 mM DTT, 0.075 μ Ci of UDP-[¹⁴C]Glc (300 mCi/mmol), and, where indicated, 0.1 mM dTDP-Rha or 0.1 mM dTDP-Rha and 0.1 mM UDP-GlcUA. Lanes show enzymes in the reaction mixtures: E, Cps2E; T, Cps2T; F, Cps2F; G, Cps2G; and I, Cps2I. Arrows indicate the locations of the origin (O) and the products formed by Cps2E, Cps2T, Cps2F, and Cps2G.

sugar of the repeat unit, we used reactions with Cps2E-, Cps2T-, Cps2F-, Cps2G-containing membranes and the nucleotide sugars UDP-[¹⁴C]Glc, dTDP-Rha, and UDP-GlcUA. Several attempts, with up to 20 μ g of total membrane protein, resulted in no new [¹⁴C]Glc-labeled product (Fig. 4 and data not shown). We also tested Cps2I activity directly by measuring the transfer of [¹⁴C]GlcUA from UDP-[¹⁴C]GlcUA to the Cps2G reaction product. In this case, we combined Cps2E-, Cps2T-, Cps2F-, Cps2G-, and Cps2I-containing membranes with UDP-Glc, dTDP-Rha, and UDP-[¹⁴C]GlcUA. Glycolipids formed in these reactions were extracted, dried, and analyzed for the incorporation of [¹⁴C]GlcUA. Membranes containing Cps2I did not incorporate [¹⁴C]GlcUA above the negative control (identical reactions replacing Cps2I membranes with control membranes). This result was in contrast to the positive control, which demonstrated [¹⁴C]GlcUA incorporation by *S. pneumoniae* membranes containing the serotype 3 synthase Cps3S (data not shown). This enzyme catalyzes the addition of alternating Glc and GlcUA residues from the respective UDP sugars in the synthesis of the serotype 3 capsular polysaccharide (43, 51).

The inability to detect a Cps2I product was not due to reduced amounts of this protein in relation to the other glycosyltransferases in the recombinant membranes, as we previously demonstrated that all of the glycosyltransferases are present at similar levels (25). We also analyzed the effect of the addition of UDP-Glc and *E. coli* cell lysates containing functional Cps2K to the reactions using independent Cps2E-, Cps2T-, Cps2F-, Cps2G-, and Cps2I-containing membranes. Cps2K is the UDP-Glc dehydrogenase necessary for synthesis of the Cps2I substrate UDP-GlcUA. These additions also did not result in detection of a Cps2I-dependent product (data not shown).

Similar to our result, others have been unable to observe the addition of terminal sugars on membrane-linked polysaccharides from *E. coli* (GlcUA) and *Klebsiella aerogenes* (GlcUA) using *in vitro* assays with membrane-associated glycosyltransferases (52, 53). These results were attributed to limitations of the *in vitro* systems. Factors that may be important in these systems include possible protein modifications or the requirement for a chaperone. An analysis of the *S. pneumoniae* phosphoproteome has identified a serine-threonine phosphorylated peptide that corresponds to Cps2I (54), suggesting that modification may be important in Cps2I function.

With only one enzyme and one linkage unaccounted for, we conclude that Cps2I adds the terminal α 1-6 GlcUA sugar to the type 2 repeat unit. Support for such a prediction comes from bioinformatic analyses, where *cps2I* is predicted to encode a retaining glycosyltransferase (36), and α 1-6 GlcUA is a retaining linkage.

Suppressor mutations in the extracytoplasmic loop reduce Cps2E activity and CPS levels. Deletions of the genes encoding glycosyltransferases (Cps2F, Cps2G, or Cps2I), the UDP-glucose dehydrogenase (Cps2K), the Wzx translocase (Cps2J), or the Wzy polymerase (Cps2H) select for isolates carrying secondary mutations, suggesting that the primary mutations are lethal in the absence of suppression (25, 34). Most of these secondary mutations (32 of 35) map to *cps2E* and are missense mutations (24 of 32) that occur in both the extracytoplasmic and cytoplasmic regions (25, 34) (Fig. 5A). Suppressor mutations in the cytoplasmic loop likely directly inhibit Cps2E Glc-1-P transferase activity, but the specific effects of suppressor mutations in the extracytoplasmic loop, as well as the function of this region, are less evident. Six strains from our collection contain different missense mutations that map to the Cps2E extracytoplasmic loop (Table 1 and Fig. 5A). Each of these mutations results in a nonconservative amino acid change affecting either residue polarity or the presence of an aromatic side chain. No capsule is detectable with either of the *S. pneumoniae* strains DJ916 (suppressor mutation H258R) and BX552 (L200F), whose primary mutations (Δ *cps2F* and Δ *cps2H*, respectively) eliminate the ability to synthesize the repeating capsule backbone and polymerize the repeat units, respectively (25, 34). However, the Δ *cps2K* mutants (BX547 [D167Y], BX605 [V196G], and BX612 [V165G]) and the Δ *cps2G* mutant (DJ921 [A162S]), each of which lacks the resulting side chain in the polymerized polymer, can synthesize high-molecular-weight polymer, albeit at very low levels. This CPS is detectable only in the membrane, perhaps because of the higher fraction of polysaccharide normally located there (25, 34). Because high levels of membrane-associated polysaccharide can accumulate in the parent without harm, we have interpreted this result to be indicative of the ability to transfer CPS from Und-P not only to peptidoglycan but also to another membrane acceptor in the parent, allowing the Und-P to be recycled (34). The low level of CPS produced by the Δ *cps2K* and Δ *cps2G* mutants may reflect a reduced ability to make this transfer, with the *cps2E* mutations resulting in less functional enzymes that slow synthesis of repeat units, thereby allowing the transfer to keep pace. The reduced transfer from Und-P could occur because of altered recognition of the incomplete repeat unit by one or more enzymes in the pathway (i.e., the Wzx transporter, the Wzy polymerase, or the as-yet-unknown enzyme that moves the polymer to its final destination).

To test the effects of the extracytoplasmic loop mutations on

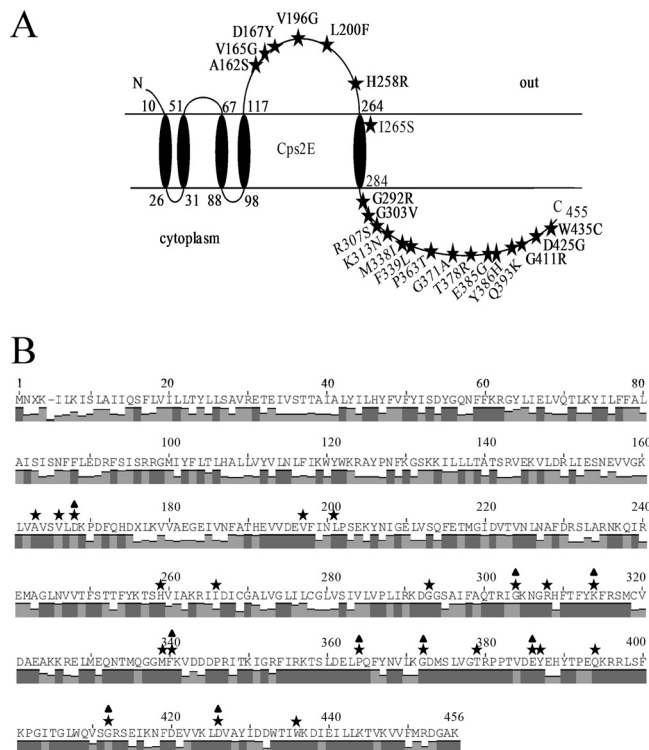


FIG 5 *S. pneumoniae* Cps2E topology and CpsE consensus sequence. (A) *S. pneumoniae* Cps2E topology and suppressor mutations (25, 34). The topology was determined using TMpred program from the ExPASy Proteomics website (<http://www.expasy.org/tools>), as previously reported (34). Only missense mutations are shown. The P363T and T378R mutations were each identified in two independent mutants. Five nonsense mutations (E191, L244, L369, W407, and Q408) and three frameshift mutations (M59, E240, and G293) have also been identified. In the original publication (34), L200F and K313N were incorrectly listed as L199F and K312N, respectively. (B) Conservation of Cps2E homologues in *S. pneumoniae* and other bacteria. Amino acid sequences from the 69 *S. pneumoniae* serotypes containing Cps2E homologues were used in a ClustalW alignment to generate a consensus sequence. Bars indicate the degree of conservation for the respective residues, where the height of the bar correlates to a greater conservation (0 to 100%; dark gray bars indicate 100% conservation). Stars above amino acids indicate relative positions of residues in the consensus CpsE where suppressor mutations were identified in the *S. pneumoniae* Cps2E. Triangles indicate residues conserved among Cps2E homologues in multiple other bacteria, including *Clostridium phytofermentans* ISDg (GenBank accession number YP_001558318), *Clostridium bolteae* (WP_002576468), *Streptococcus mitis* (WP_001013652), *Streptococcus oralis* (BAD22619), and *Streptococcus parasanguinis* (WP_003009077). Twelve residues (A177, D182, G310, G321, K331, F361, P385, G393, E407, G433, D446, and W467) are conserved in WbaP from *Salmonella enterica* subsp. *enterica* serovar 4 (YP_008266333). Note that residue numbers are shifted by one from panel A due to insertion of a gap (denoted by a hyphen) for alignment of consensus sequence. X, no conservation of residue.

Cps2E activity, we used the strains described above in *in vitro* glycosyltransferase assays. Each of the mutants exhibited activity barely above that observed for the Cps2E-negative control strain (Fig. 6A, mutations H258R, A162S, D167Y, L200F, and V196G). This result was similar to what we observed previously for $\Delta cps2K$ mutants that contained mutations in the Cps2E cytoplasmic loop (34). We again tested the latter effect using recently isolated $\Delta cps2G$ mutants containing mutations in the Cps2E cytoplasmic loop (25). Again, minimal activity was observed (M338I and R307S in Fig. 6A). In immunoblots of Cps2E, we noted reduced

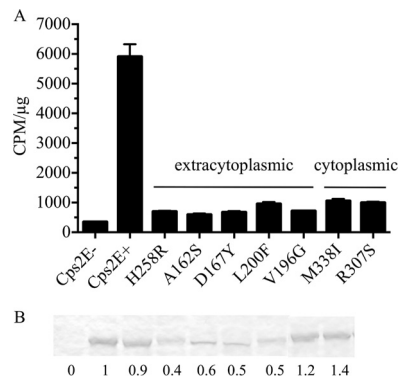


FIG 6 Cps2E glycosyltransferase activities in *S. pneumoniae* mutants containing original deletion mutations and *cps2E* suppressor mutations. (A) Membranes were used to measure incorporation of [³H]Glc from UDP-[³H]Glc to an organically soluble product. Results are the means \pm standard errors for two independent membrane preparations assayed in separate experiments (with two replicates per sample). Samples were assayed using 0.2 and 0.5 μ g of total membrane protein, which was in the linear range for the parent Cps2E activity. Activities are expressed as CPM/ μ g of total membrane protein. (B) Cps2E protein in membranes used for glycosyltransferase assays of panel A. Samples containing 5 μ g of total membrane protein were separated by SDS-12% PAGE, transferred to a nitrocellulose membrane, and probed with polyclonal antiserum to Cps2E (25). Relative protein levels were determined by densitometry and normalized to that of the Cps2E⁺ parent (D39). The values below the blot represent the means from three independent membrane preparations, two of which were used for the glycosyltransferase assays in panel A. A representative blot is shown. All samples were run on the same gel/blot; a replicate sample of M338I was spliced out of the image. Strains: Cps2E⁻, KA1522; Cps2E⁺, D39; H258R, DJ916; A162S, DJ921; D167Y, BX547; L200F, BX552; V196G, BX605; M338I, DJ918; and R307S, DJ919.

levels of protein for most of the extracytoplasmic loop mutants, but no reduction was evident for the cytoplasmic mutants in either this or the previous study (Fig. 6B and reference 34). The reduction in protein levels alone should not be sufficient to account for the severe reductions in Cps2E activity, indicating that other effects may be involved.

To address the possibility that Cps2E activities for the extracytoplasmic loop mutants might be reduced as a result of the primary deletion mutations, we introduced two of the mutations (H258R, derived from the $\Delta cps2F$ strain DJ916, and V196G, derived from the $\Delta cps2K$ strain BX605) into the parent, D39. Using membranes from these mutants, we were unable to detect Cps2E activity, even with increased amounts of membranes (Fig. 7A). As in the original deletion strains, the H258R mutant exhibited Cps2E protein levels similar to those of the parent, whereas the V196G mutant was reduced (Fig. 7B). In contrast, the non-capsule-related protein PspA, which fractionates with the membranes (44), was present in similar amounts in the parent and mutant strains (Fig. 7B). By the Quellung reaction, which uses reactivity with antiserum for microscopic observation of CPS, the mutants appeared to be CPS negative (data not shown). In the more sensitive immunoblot assay used to detect CPS in membrane and cell wall fractions, however, CPS was evident and displayed the full range of low- to high-molecular-weight polymers, albeit at reduced amounts (Fig. 7C). The amount of CPS produced was greater than that seen for the original strains containing both a *cps2E* mutation and a deletion (25), an effect we observed previously with mutations in the cytoplasmic loop (34).

To further test the effect of the extracytoplasmic loop muta-

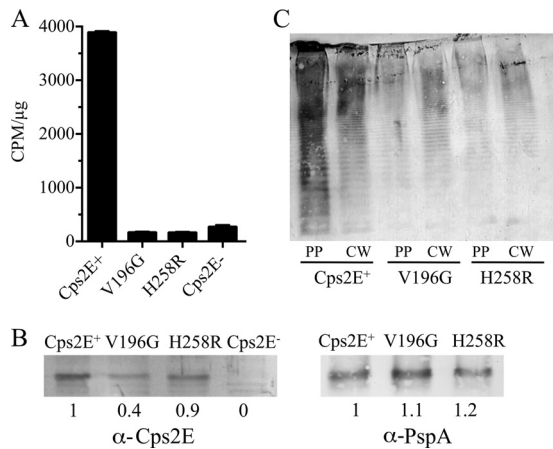


FIG 7 Cps2E glycosyltransferase activities and CPS in *S. pneumoniae* mutants containing suppressor mutations in the absence of the original deletion mutations. (A) Glycosyltransferase assays were performed as described for Fig. 6A and in Materials and Methods, except that reactions were for 2 h and contained 0.125 μCi of UDP-[^3H]Glc. For the Cps2E⁺ parent (JK903), membranes were assayed at 0.5 and 1 μg of total protein, which was in the linear range. All others were assayed at 2, 4, and 10 μg of total protein. Values are the means \pm standard errors for the two or three membrane concentrations. (B) Cps2E and PspA proteins in membranes used for glycosyltransferase assays for panel A. Experiments were performed as described for Fig. 6B, loading 2 and 4 μg of total membrane protein in separate blots. Values are the means of the two blots, normalized to the Cps2E⁺ strain. (C) CPS immunoblot of strains. Samples were separated by SDS-PAGE, transferred to nitrocellulose, and probed with polyclonal antiserum to serotype 2 CPS, as described previously (56). PP, protoplast (membrane-containing) fraction; CW, cell wall fraction. Strains: Cps2E⁺, JK903; V196G, JK906; H258R, JK907; and Cps2E⁻, KA1522.

tions on Cps2E activity, we expressed each of the mutant proteins in *E. coli*. In this case, other *S. pneumoniae* proteins that might repress activity, possibly through enhanced interactions with the mutant Cps2E proteins, would be eliminated. Once again, we were unable to detect activity (Fig. 8A). Similar to the results in *S. pneumoniae*, three of the mutant proteins (D167Y, L200F, and V196G) were detected at lower levels in the membranes than was parental Cps2E (Fig. 8B).

Two observations from this series of experiments merit further consideration. First, the production of more CPS by mutants containing only a suppressor mutation in *cps2E*, as opposed to both a suppressor mutation and the original deletion mutation, suggests that the low CPS levels in the latter strains are not due solely to reduced Cps2E activity. This observation could suggest a feedback type of regulation, where Cps2E activity is normally reduced when the end products (partial or complete repeat units) accumulate. The normal feedback inhibition may not be sufficient in the deletion mutants where excessive amounts of end products accumulate, thus necessitating suppressor mutations for survival. In the mutants containing only the *cps2E* mutations, the feedback inhibition is relieved (returned to normal levels), and CPS amounts reflect the effects of only the *cps2E* mutations. Second, the reduced levels of Cps2E in many of the mutants suggest that instability may occur as a result of the mutations. While we cannot exclude the possibility of differences in expression levels in the *E. coli* system, the observation of similar effects in *E. coli* and in *S. pneumoniae* isolates with both a deletion and a point mutation

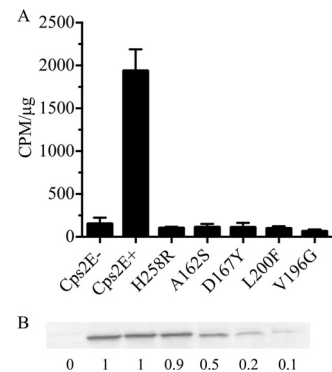


FIG 8 Cps2E glycosyltransferase activities in recombinant proteins expressed in *E. coli*. (A) Glycosyltransferase assays were performed as for Fig. 6A. Samples were assayed using 0.5 and 2 μg of total membrane protein, which was in the linear range for the Cps2E⁺ strain (KJ4152). Values are the means \pm standard errors of the two concentrations, normalized to cpm/ μg . (B) Cps2E protein in membranes used for glycosyltransferase assays for panel A. Experiments were performed as described for Fig. 6B, loading 5 μg of total membrane protein. Values are the means of two immunoblots, normalized to the Cps2E⁺ strain. Strains: Cps2E⁻, RC124; Cps2E⁺, KJ4152; H258R, DJ252; A162S, DJ253; D167Y, DJ254; L200F, DJ269; and V196G, DJ270.

or just a point mutation lends credence to this possibility. The reduced levels of Cps2E activity and CPS may result in part from this effect. For many mutants, however, the reductions in Cps2E protein levels are modest in relation to the near complete loss of *in vitro* glycosyltransferase activity. Further, reduced stability does not explain the observation of similar levels of CPS in mutants that made normal amounts of Cps2E (such as H258R and cytoplasmic loop mutants) and those that make severely reduced levels of Cps2E (other extracytoplasmic loop mutants). Perhaps this instability and reduced activity reflect a loss or severe reduction in interaction(s) with protein(s) or other factors that is more extreme *in vitro* but where interactions *in vivo* may still be sufficient for some CPS production to occur. The fact that instability occurs with mutations in the extracytoplasmic, but not cytoplasmic, loop suggests that this may be a critical region for interactions. In wild-type bacteria, such interactions may lead to enhanced Cps2E activity and CPS synthesis, but the interactions and activities are reduced by the point mutations. While interacting partners, if they exist, are not yet known, possibilities include the Wzx transporter (flippase), Wzy polymerase, Cps2A, and Cps2C, all of which have extracellular domains. Additional possibilities include the CPS polymer or repeat units and environmental factors.

Amino acids selected for suppressor mutations in Cps2E correlate with highly conserved residues across CpsE homologues in *S. pneumoniae* and other bacteria. Cps2E homologues are present in at least 69 *S. pneumoniae* serotypes (15, 45, 46). To evaluate whether residues selected as suppressor mutations in Cps2E are conserved in these 69 serotypes, we employed a ClustalW alignment to generate a consensus CpsE amino acid sequence. As indicated in Fig. 5B, 216/455 amino acids are 100% conserved across the 69 *S. pneumoniae* Cps2E homologues. To date, we have identified 22 unique missense suppressor mutations in *cps2E* (among 24 total); 17 of these affect amino acids that are 100% conserved in the 69 serotypes ($P < 0.01$) (Fig. 5B). Nine of the residues are also conserved among multiple other Gram-positive bacteria, including streptococcal and clostridial species

(Fig. 5B). Twelve of these residues (A162, D167, G292, G303, K313, F339, P363, G371, E385, G411, D425, and W435) are also conserved in their relative positions in WbaP (A177, D182, G310, G321, K331, F361, P385, G393, E407, G433, D446, and W467), the initiating glycosyltransferase for the *S. enterica* O antigen. The extracytoplasmic loops for both Cps2E and WbaP are 148 amino acids in length but share limited sequence identity (23/152 residues [15%]; positives = 49/152 [32%]; gaps = 8/152 [5%]). These results indicate that the selection for suppressor mutations in Cps2E identifies important and highly conserved functional residues.

Conclusions. The assembly of surface polysaccharides requires multiple enzymes in a complex process that is undoubtedly highly coordinated. The results obtained in this and other studies from our laboratory (24, 25, 34) and others (31) continue to suggest that Cps2E and related enzymes play central roles in the regulation of this process. The existence of the extracytoplasmic loop and the effect of mutations therein suggest a function that is necessary for interaction with other surface enzymes, repeat units/polymer, or with an environmental factor.

In *S. enterica* WbaP mutants in which only the extracytoplasmic loop is lacking, O antigen is produced at levels similar to that of the parent but with an increased frequency of polymers in the 10- to 20-repeat-unit range, and *in vitro* glycosyltransferase activity is comparable to that of the parent (31). The present results suggest that in addition to any potential role in modulation of chain length, the extracytoplasmic loop may have a role in regulating initiation of capsule synthesis. The cytoplasmic loop alone appears capable of initiating synthesis, however, as one of the extracytoplasmic loop mutations obtained in our original study contained a nonsense mutation (E191*) just prior to a secondary translation start that effectively separated the N- and C-terminal regions. This mutant was able to produce CPS, albeit at low levels (reference 34 and unpublished data). This result is consistent with the observation that the cytoplasmic region and the final transmembrane domain are sufficient for activity in other glycosyltransferases of the PHPT family (27, 31, 32, 35). In the native *S. pneumoniae* system, however, that activity may not be optimal in the absence of a fully functional outer loop.

The requirement for suppressor mutations when capsule synthesis is blocked at a point past the committed step provides a powerful tool to identify critical residues of Cps2E. Despite the fact that mutations in many other genes (including *cps2LMNO* [dTDP-Rha synthesis], *cps2C* and *cps2D* [tyrosine phosphoregulatory pathway], and *cps2T* [first rhamnosyltransferase]) could reduce or eliminate CPS synthesis, 32 of 35 suppressor mutations have thus far occurred in *cp2E*. These results suggest that there are a greater number of effective target sites in *cp2E* than in all other sites combined, even though the collective target size for the genes above is more than 4 times that of *cp2E*. We have not yet reached saturation mutagenesis; thus, many other residues likely remain to be identified, providing the potential to better understand the multiple roles of Cps2E and related enzymes in capsule synthesis.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant AI28457 from the National Institutes of Health and by the Mizutani Foundation for Glycobiology.

We thank Ella Robinson for graphics assistance.

REFERENCES

- Avery OT, Dubos R. 1931. The protective action of a specific enzyme against type III pneumococcus infection in mice. *J. Exp. Med.* 54:73–89.
- Brown EJ, Joiner KA, Cole RM, Berger M. 1983. Localization of complement component 3 on *Streptococcus pneumoniae*: anti-capsular antibody causes complement deposition on the pneumococcal capsule. *Infect. Immun.* 39:403–409.
- Griffith F. 1928. The significance of Pneumococcal types. *J. Hyg. (Lond.)* 27:113–159.
- Hardy GG, Magee AD, Ventura CL, Caimano MJ, Yother J. 2001. Essential role for cellular phosphoglucomutase in virulence of type 3 *Streptococcus pneumoniae*. *Infect. Immun.* 69:2309–2317.
- Magee AD, Yother J. 2001. Requirement for capsule in colonization by *Streptococcus pneumoniae*. *Infect. Immun.* 69:3755–3761.
- Nelson AL, Roche AM, Gould JM, Chim K, Ratner AJ, Weiser JN. 2007. Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. *Infect. Immun.* 75:83–90.
- Weiser JN, Austrian R, Sreenivasan PK, Masure HR. 1994. Phase variation in pneumococcal opacity: relationship between colonial morphology and nasopharyngeal colonization. *Infect. Immun.* 62:2582–2589.
- Wood WB, Jr, Smith MR. 1949. The inhibition of surface phagocytosis by the capsular slime layer of pneumococcus type III. *J. Exp. Med.* 90:85–96.
- Francis T, Tillett WS. 1930. Cutaneous reactions in pneumonia. The development of antibodies following the intradermal injection of type-specific polysaccharide. *J. Exp. Med.* 52:573–585.
- Macleod CM, Hodges RG, Heidelberger M, Bernhard WG. 1945. Prevention of pneumococcal pneumonia by immunization with specific capsular polysaccharides. *J. Exp. Med.* 82:445–465.
- van Dam JE, Fleer A, Snippe H. 1990. Immunogenicity and immunochemistry of *Streptococcus pneumoniae* capsular polysaccharides. *Antonie Van Leeuwenhoek* 58:1–47.
- Centers for Disease Control and Prevention. 2011. Epidemiology and prevention of vaccine-preventable diseases, 12th ed, p 233–248. Public Health Foundation, Washington, DC.
- Weiser J, Nahm MH. 2008. Immunity to extracellular bacteria, p 1182–1203. *In* Paul WE (ed), *Fundamental immunology*, 6th ed. Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia, PA.
- Abramson JS, Baker CJ, Fisher MC, Gerber MA, Meissner HC, Murray DL, Overturf GD, Prober CG, Rennels MB, Saari TN, Weiner LB, Whitley RJ. 2000. American Academy of Pediatrics. Committee on Infectious Diseases. Policy statement: recommendations for the prevention of pneumococcal infections, including the use of pneumococcal conjugate vaccine (Prevnar), pneumococcal polysaccharide vaccine, and antibiotic prophylaxis. *Pediatrics* 106:362–366.
- Bentley SD, Aanensen DM, Mavroidi A, Saunders D, Rabinowitz E, Collins M, Donohoe K, Harris D, Murphy L, Quail MA, Samuel G, Skovsted IC, Kalltoft MS, Barrell B, Reeves PR, Parkhill J, Spratt BG. 2006. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genet.* 2:e31.
- Dillard JP, Vandersea MW, Yother J. 1995. Characterization of the cassette containing genes for type 3 capsular polysaccharide biosynthesis in *Streptococcus pneumoniae*. *J. Exp. Med.* 181:973–983.
- Llull D, Garcia E, Lopez R. 2001. Tts, a processive beta-glycosyltransferase of *Streptococcus pneumoniae*, directs the synthesis of the branched type 37 capsular polysaccharide in pneumococcus and other gram-positive species. *J. Biol. Chem.* 276:21053–21061.
- Yother J. 2011. Capsules of *Streptococcus pneumoniae* and other bacteria: paradigms for polysaccharide biosynthesis and regulation. *Annu. Rev. Microbiol.* 65:563–581.
- Raetz CR, Whitfield C. 2002. Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* 71:635–700.
- Valvano MA. 2003. Export of O-specific lipopolysaccharide. *Front. Biosci.* 8:s452–s471.
- Whitfield C. 2006. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu. Rev. Biochem.* 75:39–68.
- Jansson PE, Lindberg B, Anderson M, Lindquist U, Henrichsen J. 1988. Structural studies of the capsular polysaccharide from *Streptococcus pneumoniae* type 2, a reinvestigation. *Carbohydr. Res.* 182:111–117.
- Yother J, Bentley SD, Hennessey JP. 2008. Genetics, biosynthesis, and chemistry of pneumococcal capsular polysaccharides, p 33–46. *In* Siber GR, Klugman KP, Makela PH (ed), *Pneumococcal vaccines: the impact of conjugate vaccine*. ASM Press, Washington, DC.

24. Cartee RT, Forsee WT, Bender MH, Ambrose KD, Yother J. 2005. CpsE from type 2 *Streptococcus pneumoniae* catalyzes the reversible addition of glucose-1-phosphate to a polyprenyl phosphate acceptor, initiating type 2 capsule repeat unit formation. *J. Bacteriol.* **187**:7425–7433.
25. James DBA, Yother J. 2012. Genetic and biochemical characterizations of enzymes involved in *Streptococcus pneumoniae* serotype 2 capsule synthesis demonstrate that Cps2T (WchF) catalyzes the committed step by addition of beta-1-4 rhamnose, the second sugar residue in the repeat unit. *J. Bacteriol.* **194**:6479–6489.
26. Kolkman MA, Morrison DA, Van Der Zeijst BA, Nuijten PJ. 1996. The capsule polysaccharide synthesis locus of *Streptococcus pneumoniae* serotype 14: identification of the glycosyl transferase gene *cps14E*. *J. Bacteriol.* **178**:3736–3741.
27. Pelosi L, Boumediene M, Saksouk N, Geiselmann J, Geremia RA. 2005. The glucosyl-1-phosphate transferase WchA (Cap8E) primes the capsular polysaccharide repeat unit biosynthesis of *Streptococcus pneumoniae* serotype 8. *Biochem. Biophys. Res. Commun.* **327**:857–865.
28. van Selm S, Kolkman MA, van der Zeijst BA, Zwaagstra KA, Gastra W, van Putten JP. 2002. Organization and characterization of the capsule biosynthesis locus of *Streptococcus pneumoniae* serotype 9V. *Microbiology* **148**:1747–1755.
29. Jiang XM, Neal B, Santiago F, Lee SJ, Romana LK, Reeves PR. 1991. Structure and sequence of the *rfb* (O antigen) gene cluster of *Salmonella* serovar typhimurium (strain LT2). *Mol. Microbiol.* **5**:695–713.
30. Yuasa R, Levinthal M, Nikaido H. 1969. Biosynthesis of cell wall lipopolysaccharide in mutants of *Salmonella*. V. A mutant of *Salmonella typhimurium* defective in the synthesis of cytidine diphosphoabequose. *J. Bacteriol.* **100**:433–444.
31. Saldias MS, Patel K, Marolda CL, Bittner M, Contreras I, Valvano MA. 2008. Distinct functional domains of the *Salmonella enterica* WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit. *Microbiology* **154**:440–453.
32. Wang L, Liu D, Reeves PR. 1996. C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. *J. Bacteriol.* **178**:2598–2604.
33. Wang L, Reeves PR. 1994. Involvement of the galactosyl-1-phosphate transferase encoded by the *Salmonella enterica rfbP* gene in O-antigen subunit processing. *J. Bacteriol.* **176**:4348–4356.
34. Xayarath B, Yother J. 2007. Mutations blocking side chain assembly, polymerization, or transport of a Wzy-dependent *Streptococcus pneumoniae* capsule are lethal in the absence of suppressor mutations and can affect polymer transfer to the cell wall. *J. Bacteriol.* **189**:3369–3381.
35. Patel KB, Ciepchal E, Swiezewska E, Valvano MA. 2011. The C-terminal domain of the *Salmonella enterica* WbaP (UDP-galactose:Und-P galactose-1-phosphate transferase) is sufficient for catalytic activity and specificity for undecaprenyl monophosphate. *Glycobiology* **22**:116–122.
36. Aanensen DM, Mavroidi A, Bentley SD, Reeves PR, Spratt BG. 2007. Predicted functions and linkage specificities of the products of the *Streptococcus pneumoniae* capsular biosynthetic loci. *J. Bacteriol.* **189**:7856–7876.
37. Iannelli F, Pearce BJ, Pozzi G. 1999. The type 2 capsule locus of *Streptococcus pneumoniae*. *J. Bacteriol.* **181**:2652–2654.
38. Burrows LL, Lam JS. 1999. Effect of *wzx* (*rfbX*) mutations on A-band and B-band lipopolysaccharide biosynthesis in *Pseudomonas aeruginosa* O5. *J. Bacteriol.* **181**:973–980.
39. Katzen F, Ferreira DU, Oddo CG, Ielmini MV, Becker A, Puhler A, Ielpi L. 1998. *Xanthomonas campestris* pv. *campestris* gum mutants: effects on xanthan biosynthesis and plant virulence. *J. Bacteriol.* **180**:1607–1617.
40. Tao L, LeBlanc DJ, Ferretti JJ. 1992. Novel streptococcal-integration shuttle vectors for gene cloning and inactivation. *Gene* **120**:105–110.
41. Yother J, Handsome GL, Briles DE. 1992. Truncated forms of PspA that are secreted from *Streptococcus pneumoniae* and their use in functional studies and cloning of the *pspA* gene. *J. Bacteriol.* **174**:610–618.
42. Hardy GG, Caimano MJ, Yother J. 2000. Capsule biosynthesis and basic metabolism in *Streptococcus pneumoniae* are linked through the cellular phosphoglucomutase. *J. Bacteriol.* **182**:1854–1863.
43. Cartee RT, Forsee WT, Schutzbach JS, Yother J. 2000. Mechanism of type 3 capsular polysaccharide synthesis in *Streptococcus pneumoniae*. *J. Biol. Chem.* **275**:3907–3914.
44. Yother J, White JM. 1994. Novel surface attachment mechanism of the *Streptococcus pneumoniae* protein PspA. *J. Bacteriol.* **176**:2976–2985.
45. Calix JJ, Nahm MH. 2010. A new pneumococcal serotype, 11E, has a variably inactivated *wcjE* gene. *J. Infect. Dis.* **202**:29–38.
46. Calix JJ, Porambo RJ, Brady AM, Larson TR, Yother J, Abeygunwardana C, Nahm MH. 2012. Biochemical, genetic, and serological characterization of two capsule subtypes among *Streptococcus pneumoniae* serotype 20 strains: discovery of a new pneumococcal serotype. *J. Biol. Chem.* **287**:27885–27894.
47. Troutman JM, Imperiali B. 2009. *Campylobacter jejuni* PglH is a single active site processive polymerase that utilizes product inhibition to limit sequential glycosyl transfer reactions. *Biochemistry* **48**:2807–2816.
48. Coutinho PM, Deleury E, Davies GJ, Henrissat B. 2003. An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.* **328**:307–317.
49. Ghafoor A, Jordens Z, Rehm BH. 2013. Role of PelF in pel polysaccharide biosynthesis in *Pseudomonas aeruginosa*. *Appl. Environ. Microbiol.* **79**:2968–2978.
50. Greenfield LK, Richards MR, Li J, Wakarchuk WW, Lowary TL, Whitfield C. 2012. Biosynthesis of the polymannose lipopolysaccharide O-antigens from *Escherichia coli* serotypes O8 and O9a requires a unique combination of single- and multiple-active site mannosyltransferases. *J. Biol. Chem.* **287**:35078–35091.
51. Dillard JP, Yother J. 1994. Genetic and molecular characterization of capsular polysaccharide biosynthesis in *Streptococcus pneumoniae* type 3. *Mol. Microbiol.* **12**:959–972.
52. Drummel-Smith J, Whitfield C. 1999. Gene products required for surface expression of the capsular form of the group 1 K antigen in *Escherichia coli* (O9a:K30). *Mol. Microbiol.* **31**:1321–1332.
53. Sutherland IW, Norval M. 1970. The synthesis of exopolysaccharide by *Klebsiella aerogenes* membrane preparations and the involvement of lipid intermediates. *Biochem. J.* **120**:567–576.
54. Sun X, Ge F, Xiao CL, Yin XF, Ge R, Zhang LH, He QY. 2010. Phosphoproteomic analysis reveals the multiple roles of phosphorylation in pathogenic bacterium *Streptococcus pneumoniae*. *J. Proteome Res.* **9**:275–282.
55. Yother J. 2006. Integration of capsular polysaccharide biosynthesis with metabolic and virulence pathways in *Streptococcus pneumoniae*, p 51–65. In Brogden KA, Minion FC, Cornick N, Stanton TB, Zhang Q, Nolan LK, Wannemuehler MJ (ed), *Virulence mechanisms of bacterial pathogens*, 4th ed. ASM Press, Washington, DC.
56. Bender MH, Cartee RT, Yother J. 2003. Positive correlation between tyrosine phosphorylation of CpsD and capsular polysaccharide production in *Streptococcus pneumoniae*. *J. Bacteriol.* **185**:6057–6066.
57. Avery OT, MacLeod CM, McCarty M. 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a deoxyribonucleic acid fraction isolated from pneumococcus type III. *J. Exp. Med.* **79**:137–158.
58. Briles DE, Nahm M, Schroer K, Davie J, Baker P, Kearney J, Barletta R. 1981. Antiphosphocholine antibodies found in normal mouse serum are protective against intravenous infection with type 3 *Streptococcus pneumoniae*. *J. Exp. Med.* **153**:694–705.
59. Macrina FL, Tobian JA, Jones KR, Evans RP, Clewell DB. 1982. A cloning vector able to replicate in *Escherichia coli* and *Streptococcus sanguis*. *Gene* **19**:345–353.