2013

Genetic and Biochemical Characterizations of the Capsular Glycosyltransferases in Streptococcus pneumoniae serotype 2

David Boyd Anthony James
University of Alabama at Birmingham

GENETIC AND BIOCHEMICAL CHARACTERIZATIONS OF THE CAPSULAR GLYCOSYLTRANSFERASES IN *STREPTOCOCCUS PNEUMONIAE* SEROTYPE 2

by

DAVID BOYD ANTHONY JAMES

JANET YOTHER, COMMITTEE CHAIR
KEVIN DYBVIG
MICHAEL NIEDERWEIS
DAVID PRITCHARD
CHARLES TURNBOUGH
HUI WU

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

2013

GENETIC AND BIOCHEMICAL CHARACTERIZATIONS OF THE CAPSULAR GLYCOSYLTRANSFERASES IN *STREPTOCOCCUS PNEUMONIAE* SEROTYPE 2

DAVID BOYD ANTHONY JAMES

MICROBIOLOGY

ABSTRACT

The principle objective of this dissertation is to develop a thorough understanding of the genetics and biochemistry of capsule synthesis. Specifically, it involves the identification and characterization of the capsular glycosyltransferases in *Streptococcus pneumoniae* serotype 2, with investigations into their regulation. *S. pneumoniae* (the pneumococcus) is a significant human pathogen that has the ability to persist as a commensal and then transition to a virulent state, the latter of which is a major contributor to global morbidity and mortality. The capsular polysaccharide (CPS) of *S. pneumoniae* is a protective surface structure that plays an integral role in this disease process. The vast majority of pneumococcal serotypes synthesize capsule via the Wzy-dependent pathway, which requires the coordinated activity of multiple glycosyltransferases. Apart from the initiating capsular glycosyltransferase (Cps2E for serotype 2), few reports have functionally identified or characterized the other glycosyltransferases required to complete capsule repeat unit synthesis. Here, we utilize genetic, biochemical, and bioinformatic approaches to demonstrate that Cps2T, Cps2F, Cps2G, and Cps2I are glycosyltransferases that catalyze the remaining glycosidic linkages in the *S. pneumoniae* serotype 2 capsule repeat unit. Cps2T catalyzes the second

and committed step in repeat unit biosynthesis. Cps2F is a novel rhamnosyltransferase catalyzing the third step with the sequential addition of two sugar residues. Cps2G catalyzes the fourth step, and our evidence suggests that Cps2I catalyzes the final step in repeat unit biosynthesis. An additional aim of this dissertation is to investigate whether glycosyltransferase activity may be regulated to influence capsule levels and/or distribution. Here, we generate and analyze defined mutations in the initiating capsular glycosyltransferase Cps2E and analyze the effects on capsule initiation. We identified 5 residues (A162, D167, V196, L200, H258) in the extracytoplasmic domain of Cps2E and demonstrate that these residues are highly conserved among *S. pneumoniae* serotypes and required for efficient cytoplasmic sugar transfer *in vitro*. In addition, we analyze the data generated from both aims to propose mechanisms of how glycosyltransferase activity may be regulated to influence capsule in *S. pneumoniae*.

Keywords: Bacterial genetics; Glycosyltransferases; Polysaccharide; dTDP-rhamnose biosynthesis; Suppressor mutations; Committed step

DEDICATION

I dedicate this dissertation, and any accolades that may accompany a Ph.D., to God who gifted me with the necessary wisdom and endurance toward its completion.

"Each of you should use whatever gift you have received to serve others, as faithful stewards of God's grace in its various forms. If anyone speaks, they should do so as one who speaks the very words of God. If anyone serves, they should do so with the strength God provides, so that in all things God may be praised through Jesus Christ. To him be the glory and the power for ever and ever. Amen"

-The Bible (1 Peter 4:10-11)

ACKNOWLEDGEMENTS

Many have assisted me, in various forms, toward accomplishing this goal.

I thank my dear wife Alexa for her continual love, tender care, and encouragement during these challenging years as a graduate student. I look forward to the many experiences we will encounter together along life's journey.

I thank my parents Victor and Jean, and my brother Andrew for their love and support. From youth my parents have provided a safe, nurturing environment where I learned the importance of faith, integrity, respect, hard work, and a good education. All of these have shaped me to be the man I am today. My brother has been a great friend and trustworthy companion, and for this I am grateful.

I thank Janet for the opportunity to join her research lab and for her mentorship over the years. I am grateful for all of her instructions, guidance and support toward my academic pursuits. Thanks for seeing and helping to unlock my scientific potential. I thank my thesis committee for being fair, constructive, and critical throughout the evaluation of my academic progress and supportive in obtaining a great post-doctoral position.

I thank my current and past lab mates (Aaron, Ella, Jocelyn, Kanu, Kellie, Melissa, Matt, Rob, Tom L, and Tom F) for helping to make lab life enjoyable. Thanks for the good times, scientific advice and insightful discussions.

I thank Brandon, Eneida, and Sara for their friendship; what quality times we

have had. I thank the Microbiology department and graduate community for creating a welcoming environment. Thanks to the Bedwell, Briles, Nahm, and Prevelige labs for assistance with reagents/equipments.

I thank the Olivers (Lorren, Patrice, Shardé and Djovan) for their generous hospitality, friendship and advice. Your positive influence during this time will not be forgotten.

Finally, I thank my family and friends outside of Alabama who have sent their support and encouragement during this season of my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Continued)

Figure *Page*

STREPTOCOCCUS PNEUMONIAE SEROTYPE 2 CAPSULAR GLYCOSYLTRANSFERASES: BIOCHEMICAL ACTIVITIES AND SIGNIFICANCE OF SUPPRESSOR MUTATIONS

INTRODUCTION

Key characteristics of *Streptococcus pneumoniae*

Streptococcus pneumoniae is a remarkable bacterium by virtue of its significant impact on human health, and the transformative knowledge gained from its study. *S. pneumoniae* is the etiological agent of numerous human diseases that cause global morbidity and mortality (69). Since its initial discovery over 130 years ago, researchers studying this bacterium have unraveled many essential principles which are today at the very core of biology. Among these discoveries is the identification of DNA as the material of genetic information (5), the development of the Gram stain (29), the discovery of bacterial quorum sensing (79), and the immunogenicity of polysaccharides (75) - leading to the beginning of quantitative immunology (34).

S. pneumoniae has been known by many names, including *Microbe septicemique du salive* (64), *Micrococcus pasteuri* (77), and *Diplococcus pneumoniae* (95) before arriving at the current scientific name in 1974. Today, it is commonly referred to as the pneumococcus. Regardless of the name, its biological properties and global influence as a human pathogen remain the same. It is characterized as a Gram-positive, α-hemolytic, often encapsulated, bile soluble, aerotolerant, lancet shaped bacterium.

Pathogenesis

The natural niche for *S. pneumoniae* is the human nasopharynx. Colonization, or carriage, in the nasopharynx is the absolute and first requirement for disease. Colonization, however, but does not necessitate active disease. In fact, reports indicate

that 20-50% of healthy individuals are asymptomatically colonized by *S. pneumoniae* (69). Transmission of pneumococci occurs through aspiration of respiratory droplets from infected individuals. Disease occurs under conditions that compromise host health, where *S. pneumoniae* can disseminate from the nasopharynx and proliferate in normally sterile areas.

Over one million deaths occur globally each year from pneumococcal diseases (69). The majority of these occur in children under the age of 5, the elderly, and those with underlying immune deficiencies such as HIV-infected individuals (69). A common pneumococcal infection among children is otitis media (middle ear infection). Here, *S. pneumoniae* travels from the nasopharynx through the eustachian tube and proliferates in the middle ear. Among the elderly and immunocompromised, *S. pneumoniae* commonly bypasses protective immune responses and proliferates in the lungs, (causing pneumonia), blood (bactermia), or brain (meningitis). *S. pneumoniae* is the leading cause of community acquired pneumonia and the leading cause of pediatric meningitis in the United States (69).

The most important virulence factor of *S. pneumoniae* is the capsular polysaccharide, the synthesis of which is discussed in detail later. The capsule is a protective surface structure that surrounds the bacterium (See Fig. 1). *S. pneumoniae* mutants deficient in capsule production, relative to their capsule-producing parental strains, poorly colonize the nasopharynx and are avirulent in animal models of pneumococcal disease (4, 30, 32, 54). The mechanism of protection afforded by the capsule is to inhibit complement-mediated phagocytosis (discussed more below). There are over 90 different types of capsule, each of which is serologically distinct (10, 14, 35,

81). These capsule types, or serotypes, are differentiated by their carbohydrate composition, glycosidic bonds between sugar residues, and modification of sugars by small molecules. Serotypes can be grouped by their similarity into serogroups.

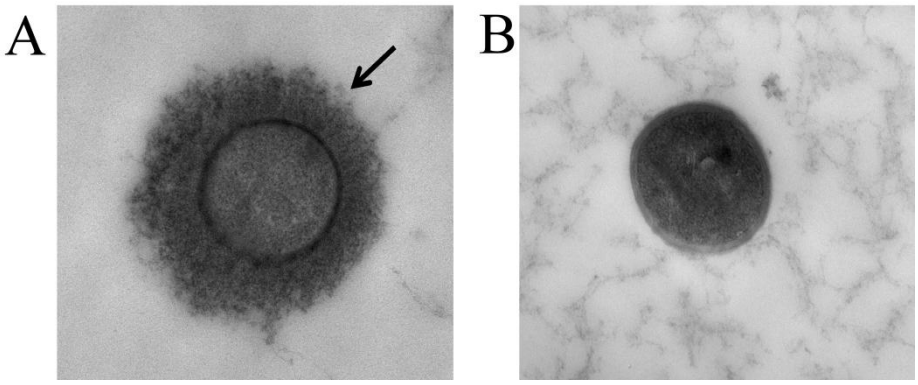

FIG 1. Electron micrographs of *S. pneumoniae*. (A) *S. pneumoniae* D39, an encapsulated strain expressing type 2 capsule. (B) *S. pneumoniae* AM1000, a non-encapsulated isogenic *cps2A-I* mutant of D39. The arrow indicates the capsule. Pictures were taken and kindly provided by Kanupryia Gupta.

The polysaccharide capsule in colonization and virulence

Capsule levels are not static but dynamic, and depend on the stage of the infection. During nasopharyngeal colonization, lower levels of capsule are required for sufficient passage through the mucosal layer and binding to epithelial cells (47, 54, 61, 91). The reduction of capsule also facilitates a greater exposure to previously masked surface adhesions (e.g., teichoic acids) and enzymes that degrade host polysaccharides (e.g., exoglycosidases). During systemic infections, capsule levels are high and provide the necessary protection against rapid host clearance (4, 32, 52, 54).

This shift in the display of pneumococcal surface structures is further exemplified by the observation of transparent and opaque phenotypes. The transparent phenotype is

characterized as smaller colonies on agar plates, reduced capsule levels, and increased teichoic acids (47, 91). The opaque phenotype is evident by larger colonies on agar plates, increased capsule, and reduced teichoic acids (47, 91). Transparent variants are commonly isolated from the nasopharynx while opaque variants are commonly isolated from systemic sites such as the blood. The necessity to modulate capsule levels during different stages of infection indicates that this regulation is a critical process for pneumococcal pathogenesis.

Serogroups also differ in their relative efficiency to colonize and cause disease. A recent review of multiple epidemiological studies (>70) concluded that ten serogroups frequently caused pediatric invasive diseases in their respective geographic regions (33). Among these, serogroups 1, 6, 14, 19 and 23 were the most prevalent. The mechanistic basis for these differences is not completely understood, but capsule switching experiments have revealed the importance of the genetic background in addition to the expressed capsule (45); indicating that the differences among serogroups is the accumulation of multiple genotypic and phenotypic differences.

Immune response

Both arms of immunity (innate and adaptive) play important roles in host protection against *S. pneumoniae* diseases. The innate immune system (inflammation, neutrophil influx, complement deposition) serves as the primary method of detection and clearance of invading microbes, while the adaptive immune system (B- and T-cell activation) provides a lasting memory-based method of quickly clearing previously

encountered pathogens (90). The pneumococcus, however, has developed methods to evade the innate immune system and limit the effects of the adaptive immune system.

Initial binding of pneumococci to the respiratory epithelium results in a mild acute inflammatory response and the corresponding infiltration of neutrophils (55, 100). This neutrophil influx, however, is insufficient to eradicate pneumococcal colonization due to the presence of capsule. As introduced earlier, the capsule serves as a protective structure that limits recognition of bound complement and, to a lesser degree, complement deposition (2, 12, 36). Complement represents a diverse set of compounds within the innate immune system that function to support antibody mediated clearance by acting as opsonins (85, 86). Thus, the capsule protects against complement-mediated opsonophagocytosis. *S. pneumoniae* produces a surface protease (Iga) that also functions to inhibit antibody receptor-mediated phagocytosis by cleaving the tail region (Fc domain) from immunoglobin A1 (IgA1) (46). Other virulence factors, including the complement inhibiting pneumococcal surface protein A (PspA), reviewed in references 40 and 42 contribute to the ability of pneumococci to evade the innate immune system.

To resolve nasopharyngeal colonization, a gradual monocyte/macrophage infiltrate is required to mount an adaptive immune response (100). Here, antibodies are ultimately generated to a variety of pneumococcal surface structures including the polysaccharide capsule. Anti-capsule antibodies are specific to the infecting serotype (53) and function to enhance opsonophagocytic killing. The few colonizing pneumococci that survive this response presumably have incorporated new genomic DNA that results in formation of new a capsule structure, a potential mechanism that limits the effects of antibody mediated clearance. Support for this conjecture comes from studies where

clinical isolates expressing varying capsule types were found to represent the same genetic background (6, 22, 23, 62); a serological variation that is the result of recombination events within the capsule locus and not spontaneous mutations (24).

Therapeutic interventions

Early studies with *S. pneumoniae* identified that the capsule is immunogenic (28), and antibodies generated against a particular capsule type were protective against pneumococcal disease (53). Consequently, the development of a safe and effective pneumococcal polysaccharide vaccine (PPV) began; and in 1977 the first vaccine was licensed in the United States. The most successful polysaccharide vaccine, PPV23, was licensed in 1983 and modern versions are currently in use today. PPV23 contains 23 different capsule types that represent serotypes that account for 88% of pneumococcal bacteremia (69).

In adults, the polysaccharide antigens in PPV23 mount an effective B cell response to produce protective antibodies and immunological memory (90). Children under the age of two, however, have immature B cells and polysaccharides are poorly immunogenic (90). In 2000 a new vaccine, PCV7, was licensed in the United States that conjugated a *Corynebacterium diphtheriae* toxoid to purified capsule from 7 serotypes (69). This conjugation activates T cells in children under the age of two and provides a sufficient immunological response to produce protective antibodies and immunological memory (90). PCV7 represented the 7 most common serotypes that cause 80% of pneumococcal infection in children younger than six years (69). PCV7 was then expanded in 2010 to PCV13 which provided protection from six additional serotypes.

Since implementation of these vaccines, significant reductions in the overall burdens of pneumococcal diseases have been observed. However, limitations on the efficacy of these vaccines to prevent pneumococcal disease still remain; most notably the coverage of only select serotypes. The advent of these vaccines, particularly the conjugate vaccines, has resulted in a phenomenon known as serotype replacement where serotypes not included in the vaccine now occupy the vacated niche (21). Serotype replacement, along with the mounting observations of multidrug resistant clinical *S. pneumoniae* isolates (69, 78, 94), underscores the importance of identifying better therapeutics and therapeutic targets that are not serotype-dependent.

Mechanisms of capsule synthesis

Three pathways for synthesizing capsules have been identified in bacteria: Wzy-, ABC- and synthase- dependent mechanism (70, 80, 93). These pathways are based on their similarity to lipopolysaccharide O-antigen biosynthetic mechanisms.

The ABC (ATP-Binding Cassette)-dependent pathway is utilized only by Gram-negative bacteria and facilitates the export of polysaccharides at the expense of ATP hydrolysis (93). Here, glycosyltransferases in the cytoplasm sequentially build the capsule polymer on a membrane acceptor which is closely associated with a multi-enzyme transport complex comprised of two domains: the nucleotide binding domain (NBD) and the transmembrane domain (TMD). The NBD is located on the cytoplasmic face of the inner membrane and, via ATP hydrolysis, facilitates the transport of the growing polymer to the TMD. The TMD is comprised of integral membrane proteins in the inner and outer membrane that facilitate physical passage of the capsule polymer

from the cytoplasm to the outside surface of the outer membrane.

In contrast to the ABC-dependent pathway, both Gram-positive and Gram-negative bacteria can utilize either the Wzy- or synthase pathway (70, 80, 97). The synthase pathway involves an integral membrane protein that catalyzes the formation, polymerization, and transport of the polysaccharide across the plasma membrane (18, 27, 70). Synthases belong to the protein family of glycosaminoglycan glycosyltransferases and are also present in eukaryotes (26). Glycosaminoglycans represent a class of unbranched polysaccharides composed of repeating disaccharide units containing a hexosamine. In the Wzy-pathway (described below in greater detail for *S. pneumoniae* type 2 capsule synthesis), multiple enzymes sequentially build the polysaccharide repeat unit which is ultimately polymerized by an enzyme belonging to the Wzy family of polymerases (70, 80, 97).

The *S. pneumoniae* polysaccharide capsule

S. pneumoniae serotypes utilize either the Wzy- or synthase dependent pathway to synthesize capsule. All but two of the 90-plus serotypes utilize the Wzy-dependent pathway while the remaining, serotypes 3 and 37, utilize the synthase-dependent pathway. This study focuses on the Wzy-dependent pathway and all further discussion will be based on this mechanism.

As a model for studying the Wzy-dependent pathway in *S. pneumoniae,* our laboratory uses serotype 2. The type 2 capsule is a repeating, branched hexasaccharide structure comprised of three different sugar residues (two glucose [Glc], three rhamnose [Rha] and one glucuronic acid [GlcUA] per repeat unit) (17, 39) (Fig. 2). Our laboratory

has previously demonstrated that serotype 2 capsule assembly begins with the initiating glycosyltransferase Cps2E (17). This glycosyltransferase, described in greater detail below, transfers glucose-1-phosphate (Glc-1-P) from a nucleotide sugar donor to a lipid acceptor (undecaprenol phosphate, Und-P) in the membrane. The subsequent steps of capsule synthesis involve additional glycosyltransferases, predicted to be Cps2T, Cps2F, Cps2G, and Cps2I, that sequentially complete repeat unit synthesis on Und-P (1, 37). The complete repeat unit is then enzymatically translocated to the outside surface of the cytoplasmic membrane (Wzx flippase) and polymerized (Wzy polymerase). The capsule polymer is then released, attached to the cell wall, or remains associated with the membrane.

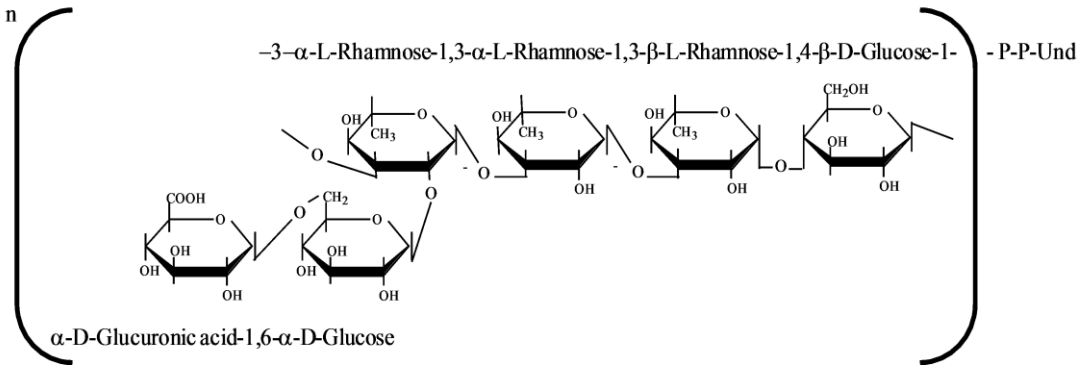

FIG 2. Haworth projection of *S. pneumoniae* type 2 capsule repeating unit linked to Undecaprenol pyrophosphate.

Genetics of the capsule locus

The capsule locus of *S. pneumoniae* encodes all the enzymes required for assembling the repeat unit (Fig. 3). For all Wzy-dependent serotypes, this genetic region is organized in a similar cassette-like arrangement. Here, the capsule locus is situated in the same region of the chromosome between the genes *dexB* (encoding dextran glucosidase B) and *aliA* (Ami-like ABC transporter permease). The natural competence of *S. pneumoniae* facilitates frequent genetic exchange of serotypes via homologous recombination between the conserved flanking regions.

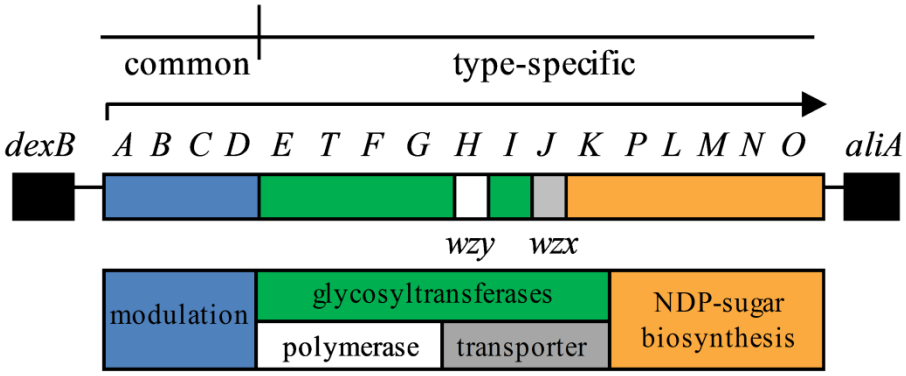

FIG 3. Genetic locus for the serotype 2 capsule. The arrow indicates the direction of transcription and the predicted operon.

The capsule loci for these serotypes can be divided into two regions, a 5' common region that is conserved across all Wzy serotypes and a 3' type-specific region. The common region encodes CpsABCD, which are involved in modulation of capsule levels. CpsA is a putative regulator of capsule levels (58), although its specific mechanism of action remains to be reported. CpsC and CpsD encodes two components of an autophosphorylating tryrosine kinase, while CpsB serves as the cognate tyrosine

phosphatase (9, 57). Downstream phosphorylation targets, and how this modification influences capsule levels, remains to be identified. The type-specific region varies for all serotypes and encodes the glycosyltransferases, flippase, polymerase and enzymes for donor sugar biosynthesis specific to the serotype. The capsule locus is predicted to be transcribed as a single polycistron from a promoter upstream of *cpsA* (37).

Glycosyltransferase function, folds and classification

Glycosyltransferases represent one of the most important classes of enzymes on the planet. Present in every kingdom of life, these enzymes catalyze the transfer of a carbohydrate moiety from a donor substrate to varied acceptors; including other saccharides, proteins, lipids, DNA and numerous small molecules (11, 51, 89). Donor substrates for these enzymatic reactions are usually activated in the form of nucleoside diphosphate sugars (11, 15, 16, 51, 89), but other donors such as lipid phosphates and nucleoside monophosphates have also been identified (89).

Two major folds, GT-A and GT-B, have been identified in all solved crystal structures of glycosyltransferases (11, 51, 89). Both the GT-A and GT-B folds contain two characteristic Rossmann-like β/α/β domains. The Rossmann fold is a structural motif commonly found in proteins that bind nucleotides (71). The differentiating factor between the GT-A and GT-B folds is that the Rossmann domains are more loosely associated in the GT-B fold. Glycosyltransferases possessing transmembrane domains adopt a different classification and can be grouped into families based on their topology and the sugar transferred (80). One such example includes the integral membrane glycosyltransferase Cps2E and its classification as a polyprenyl-phosphate hexose-1-

phosphate transferase (PHPT) (discussed further below in the Cps2E functional domains section.)

Nucleotide sugar donor biosynthesis and utilization

Donor substrates for glycosyltransferases are derived from the glycolysis intermediate glucose-6-phosphate and commonly converted to nucleotide sugars (Fig. 4). The most common nucleotide sugar utilized among *S. pneumoniae* serotypes is uridine diphosphate alpha-D-galactose (UDP-Gal), followed by uridine diphosphate alpha-D-glucose (UDP-Glc) and then deoxy thymidine diphosphate beta-L-Rha (dTDP-Rha) (10).

Glycosyltransferases utilizing nucleotide sugar donors transfer the sugar moiety via an inverting or retaining mechanism. In these mechanisms, the anomeric configuration of the donor sugar is either retained or inverted as it is transferred to the acceptor (11, 51, 89). For example, an inverting glycosyltransferases utilizing dTDP-β-Rha will form an α-Rha linkage in the reaction product.

***S. pneumoniae* capsular glycosyltransferases**

With regard to the Wzy-dependent mechanism for capsule synthesis, there are a limited number of publications characterizing pneumococcal capsular glycosyltransferases, and even less on how they are regulated. The initiating glycosyltransferase CpsE has been the most characterized, with analysis of enzyme function demonstrated in serotypes 2, 8, 9v and 14 (17, 48, 66, 82).

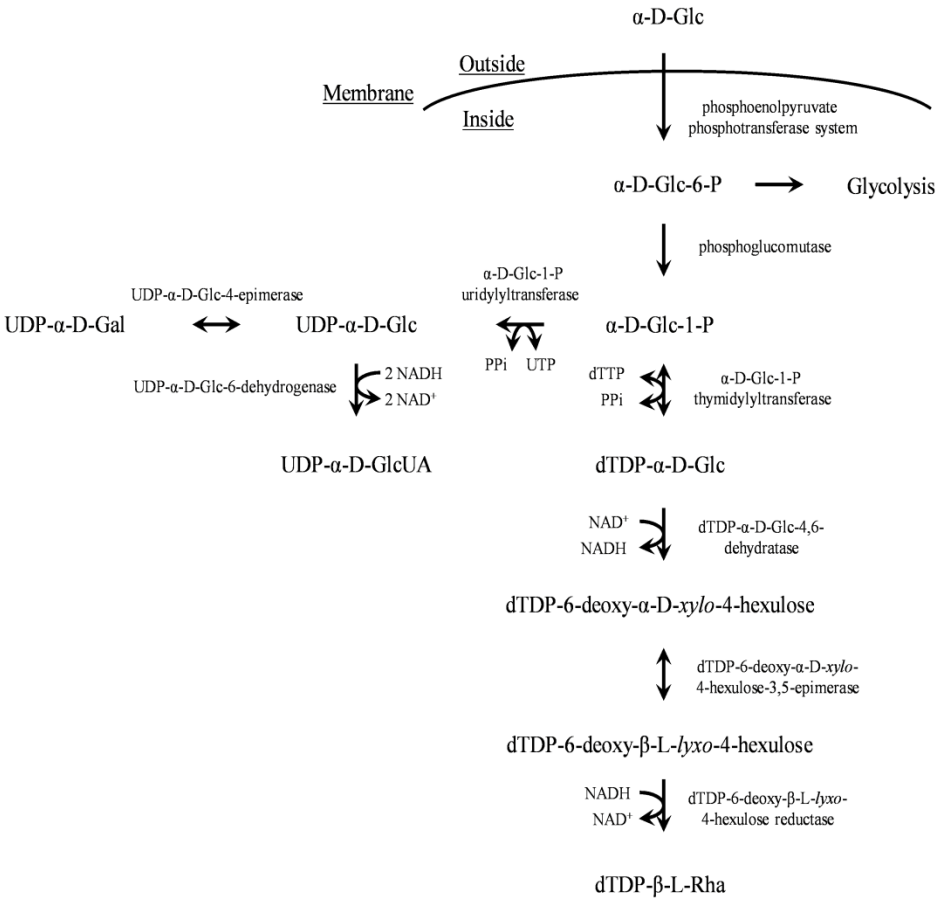

FIG 4. Nucleotide sugar biosynthesis in *S. pneumoniae*. Exogenous Glc is transported into the cell and phosphorylated via the phosphoenolpyruvate phosphotransferase system, generating Glc-6-P (68). Glc-6-P is converted to the respective nucleotide sugars by the indicated enzymes.

Here, similar to serotype 2, CpsE transfers Glc-1-P from uridine diphosphate glucose (UDP-Glc) to Und-P. These four serotypes, however, differ vastly in the composition of subsequent sugar residues added (10). Pneumococcal glycosyltransferase activities, involved in the subsequent steps of capsule repeat unit formation, have been demonstrated for glucose (Glc), galactose (Gal), and N-acetylglucosamine (GlcNAc) transfer (49, 50, 73) . However, glycosyltransferases that transfer GlcUA or Rha still remain to be experimentally identified and characterized in *S. pneumoniae*. GlcUA and

Rha are present in approximately 20% and 44%, respectively, of the published *S. pneumoniae* capsular structures (10).

In serotype 2, four open reading frames within the capsule locus (*cps2T*, *cps2F*, *cps2G* and *cps2I*) are predicted to encode glycosyltransferases that function after Cps2E (1, 10, 37). This bioinformatic prediction is primarily based on a similarity to conserved GT-A and GT-B folds that are consistent with most glycosyltransferases. In addition, the primary amino acid sequences of Cps2T, Cps2F, Cps2G and Cps2I are comparable to other glycosyltransferases, which facilitates their classification into one of the 90-plus glycosyltransferase families (Table 1) (1, 10).

TABLE 1. Glycosyltransferases encoded in the *S. pneumoniae cps2* locus

Gene	Predicted GT fold, GT family and mechanism of catalysis [a]
cps2E	None, PHPT family (Inverting)
cps2T	GT-B, GT4 (Retaining)
cps2F	GT-A, GT2 (Inverting)
cps2G	GT-B, GT4 (Retaining)
cps2I	GT-B, GT4 (Retaining)

[a] Predicted GT fold, family and mechanism of catalysis is based on (1). Cps2E represents a class of integral membrane enzymes that do not have solved crystal structures and are classified based on their topology and the sugar transferred (80).

Cps2E functional domains

Cps2E is the initiating capsular glycosyltransferase that transfers Glc-1-P from UDP-Glc to Und-P, generating Und-P-P-Glc (17). Cps2E is a member of the polyprenyl-phosphate hexose-1-phosphate transferase (PHPT) family (80). The most well-characterized PHPT enzyme is *Salmonella enterica* galactose-1-phosphate transferase WbaP, which initiates O-antigen biosynthesis (41, 98). Like other PHPT family members (80), Cps2E activity is readily reversible *in vitro* (17).

Studies of WbaP have revealed three functional domains within PHPT enzymes: the N-terminal domain, comprising 4 membrane spanning regions; a central domain, containing an extracytoplasmic loop; and the C-terminal domain, comprising 1 membrane spanning region and a cytoplasmic tail (74). The C-terminal domain catalyzes sugar transfer, as this domain is alone sufficient for glycosyltransferase activity in WbaP (65, 74, 87). The N-terminal domain is suggested to be required for protein insertion/stability in the membrane (74) and/or involved in release of Und-P-P-linked sugars from the protein (87). The function of the extracytoplasmic loop is less clear. This domain is speculated to play a role in modulation of the polysaccharide chain-length distribution (74), although the exact mechanism of action remains to be determined.

Among the 90-plus *S. pneumoniae* serotypes, at least 69 contain homologues of Cps2E (also known as WchA) (10). Each of these serotypes initiates repeat unit synthesis with the addition of glucose. Other *S. pneumoniae* initiating capsular glycosyltransferases begin repeat unit synthesis with other hexoses (primarily galactose) but lack the extracytoplasmic loop. The basis for the association between the extracytoplasmic loop

and Glc transfer in *S. pneumoniae* remains to be understood; *S. enterica* WbaP contains an extracytoplasmic loop but transfers galactose.

Alteration in capsule assembly and spontaneous suppressor mutations

Capsule assembly is multifaceted, requiring several enzymes to sequentially coordinate repeat unit synthesis. In *S. pneumoniae* serotype 2, assembly occurs on a membrane lipid acceptor with properties consistent with undecaprenyl phosphate (Und-P) (17). Und-P is utilized by other cellular processes that are essential for viability, such as peptidoglycan synthesis. Levels of Und-P in both Gram-positive and Gram-negative bacteria are extremely low (7, 17, 56). As a result, Und-P utilized in the capsule pathway must be recycled in order to maintain sufficient levels for viability.

Our laboratory has previously observed that deletion of specific genes within the *cps2* locus not only stalls capsule synthesis but results in cell death. These genes include *cps2K* (encoding UDP-Glc dehydrogenase), *cps2J* (flippase), and *cps2H* (polymerase) (96). Mutation of these genes results in either partial repeat unit formation due to lack of the terminal GlcUA, inhibition of repeat unit transport, or inhibition of repeat unit polymerization. The lethality of these mutations is hypothesized to be due to 1) failure to transfer the repeat unit or polymer from Und-P to another acceptor, resulting in sequestration of Und-P in the Wzy pathway or 2) destabilization of the membrane by accumulation of Und-P-P-linked repeat units (96).

The only way these mutants can be isolated is if a spontaneous secondary mutation occurs in the transformant that suppresses, or functionally compensates for, the primary mutation. For *S. pneumoniae* type 2, these secondary suppressor mutations

consistently mapped to genes within the capsule locus, primarily *cps2E* (96). Mutations in *cps2E* would function to suppress the lethality of the primary mutation by eliminating or significantly reducing the initiation of capsule synthesis.

AIMS OF DISSERTATION

The primary aim of this dissertation is to identify and characterize the capsular glycosyltransferases in *S. pneumoniae* serotype 2. The vast majority of pneumococcal serotypes synthesize capsule via the Wzy-dependent pathway, which requires the coordinated activity of multiple glycosyltransferases. Apart from the initiating capsular glycosyltransferase (Cps2E for serotype 2), few reports have functionally identified or characterized the other glycosyltransferases required to complete capsule repeat unit synthesis. Here, we utilize genetic and biochemical approaches to identify and characterize the remaining capsular glycosyltransferases in *S. pneumoniae* serotype 2. These glycosyltransferases are predicted to be Cps2T, Cps2F, Cps2G, and Cps2I. Determining the functions of these enzymes also reveals their reaction order in repeat unit synthesis, which provides a framework to understand how these glycosyltransferases might interact to regulate capsule synthesis. An additional aim of this dissertation is to investigate whether glycosyltransferase activity may be regulated to influence capsule levels and/or distribution. Here, we generate and analyze defined mutations in the initiating capsular glycosyltransferase Cps2E and analyze the effects on capsule. In addition, we analyze the data generated from both aims to propose mechanisms of how glycosyltransferase activity may be regulated to influence capsule production in *S. pneumoniae*.

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 β1-4 RHAMNOSE, THE SECOND SUGAR RESIDUE IN THE REPEAT UNIT

DAVID B. A. JAMES AND JANET YOTHER

Journal of Bacteriology
[Dec; 194(23): 6479-89. doi: 10.1128/JB.01135-12]

ABSTRACT

Five genes (*cps2E*, *cps2T*, *cps2F*, *cps2G*, *cps2I*) are predicted to encode the glycosyltransferases responsible for synthesis of the *Streptococcus pneumoniae* serotype 2 capsule repeat unit, which is polymerized to yield a branched surface structure containing glucose-glucuronic acid linked to a glucose-rhamnose-rhamnose-rhamnose backbone. Cps2E is the initiating glycosyltransferase, but experimental evidence supporting the functions of the remaining glycosyltransferases is lacking. To biochemically characterize the glycosyltransferases, the donor substrate dTDP-rhamnose was first synthesized using recombinant *S. pneumoniae* enzymes Cps2L, Cps2M, Cps2N, and Cps2O. In *in vitro* assays with each of the glycosyltransferases, only reaction mixtures containing recombinant Cps2T, dTDP-rhamnose, and the Cps2E product (undecaprenyl pyrophosphate glucose) generated a new product, which was consistent with lipid-linked glucose-rhamnose. *cps2T*, *cps2F*, and *cps2I* deletion mutants produced no detectable capsule, but trace amounts of capsule were detectable in Δ*cps2G* mutants, suggesting that Cps2G adds a nonbackbone sugar. All Δ*cps2F*, Δ*cps2G*, and Δ*cps2I* mutants contained different secondary suppressor mutations in *cps2E*, indicating that the initial mutations were lethal in the absence of reduced repeat unit synthesis. Δ*cps2T* mutants did not contain secondary mutations affecting capsule synthesis. The requirement for secondary mutations in mutants lacking Cps2F, Cps2G, and Cps2I indicates that these activities occur downstream of the committed step in capsule synthesis and reveal that Cps2T catalyzes this step. Therefore, Cps2T is the β1-4 rhamnosyltransferase that adds the second sugar to the repeat unit and, as the committed

step in type 2 repeat unit synthesis, is predicted to be an important point of capsule regulation.

INTRODUCTION

Streptococcus pneumoniae (the pneumococcus) is a significant human pathogen (13) that generally persists as a commensal colonizer of the nasopharynx but has the capacity to transition to a virulent state, potentially causing pneumonia, meningitis, or bacteremia. The capsular polysaccharide (CPS) plays an integral role in this process. High levels of CPS serve as a barrier against host complement-mediated opsonophagocytosis and limit access to underlying bacterial surface components during infection, but only low levels are required to reduce clearance during nasopharyngeal colonization (2, 7, 24, 28, 38, 44, 57, 60). *S. pneumoniae* strains that are deficient in CPS production are avirulent (2, 24, 28, 38), indicating that the capsule, its synthesis, and regulation are all critical processes for pneumococcal pathogenesis.

Three pathways for synthesizing CPS in bacteria have been identified and are based on their similarities to the lipopolysaccharide O-antigen biosynthetic mechanisms: the Wzy-, ABC transporter-, and synthase-dependent pathways (48, 52, 59, 62). This study focuses on the Wzy-dependent pathway, which is the predominant mechanism for CPS synthesis in Gram-positive bacteria. Here, the capsule repeat unit is systematically synthesized in the cytoplasm on a membrane lipid acceptor (undecaprenyl phosphate [Und-P]) by multiple glycosyltransferases. Once the repeat unit is formed, it is enzymatically transferred across the cytoplasmic membrane by the Wzx flippase and

polymerized by the Wzy polymerase (Fig. 1). In addition to CPS and O antigens, the Wzy dependent pathway is also used to synthesize many exopolysaccharides (25). Characterization of the enzymes in this pathway may thus reveal biosynthetic and regulatory properties that are conserved among polysaccharides produced by bacteria of medical and industrial importance.

Over 90 distinct capsular serotypes have been reported for *S. pneumoniae*, each differentiated by their polysaccharide composition and/or glycosidic bonds (5, 9, 29, 53). Two of these serotypes synthesize capsule via the synthase-dependent pathway, whereas the remaining serotypes utilize the Wzy-dependent pathway (5, 15, 37, 62). As a model for studying the Wzy-dependent pathway in *S. pneumoniae*, our laboratory uses serotype 2 (Fig. 1). Previous work in our laboratory characterized the initiating glycosyltransferase, Cps2E (12). This integral membrane enzyme transfers glucose-1-phosphate (Glc-1-P) to a membrane-bound polyprenyl phosphate acceptor with properties consistent with those of Und-P (12). In addition to *cps2E*, 4 other open reading frames within the type 2 capsule locus (*cps2T*, *cps2F*, *cps2G*, and *cps2I*) are predicted to encode glycosyltransferases (1, 32, 39), but biochemical evidence supporting their functions is lacking.

Donor substrates for glycosyltransferase reactions are usually activated in the form of nucleoside diphosphate sugars (6, 10, 11, 36, 56). dTDP rhamnose (dTDP-Rha) is one of the nucleotide-sugar donors utilized by *S. pneumoniae* capsular glycosyltransferases. dTDP-Rha is not commercially available but can be synthesized from Glc-1-P via four enzymatic reactions: α-D-Glc-1-P → dTDP-α-D-Glc → dTDP–6-deoxy-α-D-*xylo*-4-hexulose → dTDP-6-deoxy-β-L-*lyxo*-4-hexulose → dTDP-β-LRha

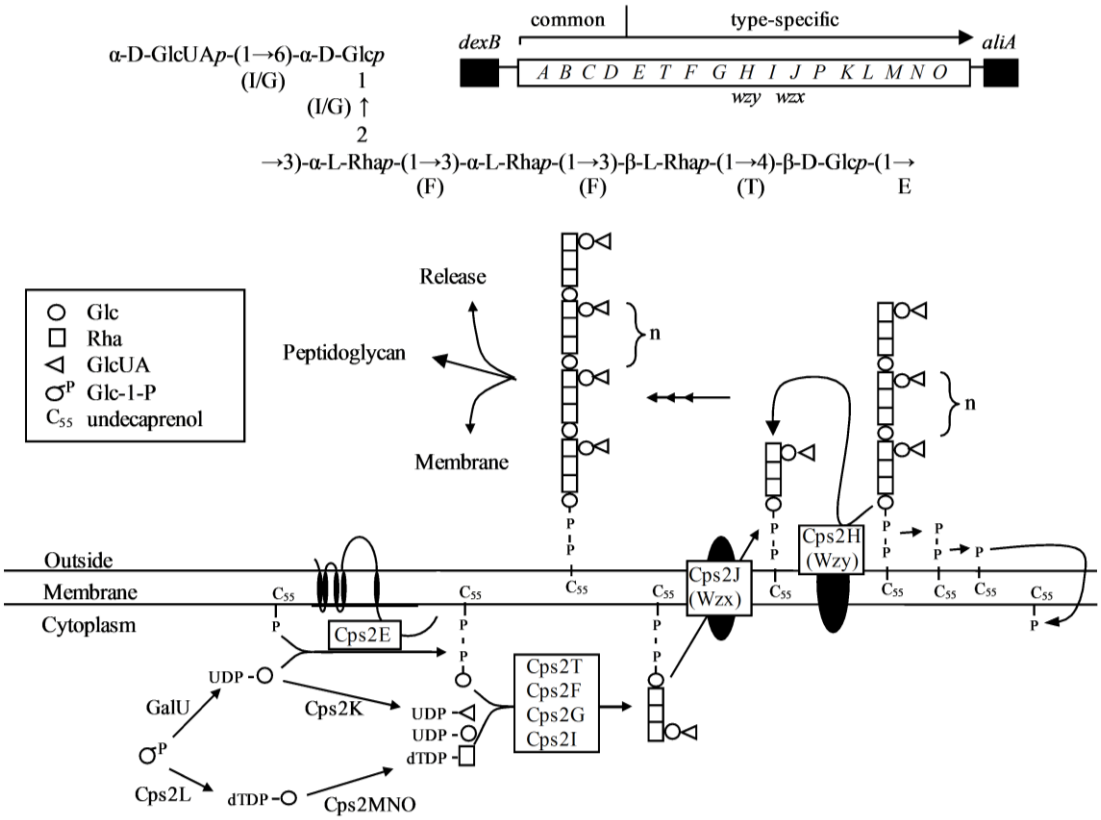

FIG 1. *S. pneumoniae* type 2 genetic locus, capsular polysaccharide repeat unit structure, and Wzy-dependent pathway. (Top) *cps2* genetic locus. The arrow indicates the direction of transcription and the predicted operon. (Middle) Repeat unit structure. Letters below each glycosidic linkage represent the glycosyltransferase known or predicted to catalyze the linkage. Predicted enzymes are in parenthesis and are based on (1), where the specificities of Cps2G and Cps2I were unclear. Glc, glucose; Rha, rhamnose; GlcUA, glucuronic acid. E, Cps2E; T, Cps2T; F, Cps2F; G, Cps2G; I, Cps2I. (Bottom) Wzy-dependent model for synthesis of the type 2 repeat unit. Synthesis of the capsule repeat unit occurs through the activity of multiple glycosyltransferases. The completed repeat unit is transferred to the outer face of the cytoplasmic membrane, polymerized and then remains associated with the membrane, released or attached to peptidoglycan. With the exception of GalU (the UDP-glucose pyrophosphorylase GalU, encoded by *galU*) all gene products are encoded in the *cps2* locus. Adapted from reference (63). Used with permission of ASM Press.

(20-22, 35, 46). These reactions are predicted to be catalyzed by 4 enzymes encoded within the type 2 locus: Cps2L, Cps2N, Cps2M, and Cps2O, respectively (32). *cpsLMNO* from serotype 19 encode enzymes that can functionally complement an *Escherichia coli* strain where the homologous dTDP-Rha biosynthesis genes, *rfbBDAC*, are deleted (42). *In vitro* activities and synthesis of dTDP-Rha, however, have not been demonstrated for the serotype 19 enzymes. Biochemical analyses of dTDP-Rha biosynthesis enzymes from *Pseudomonas aeruginosa* and *Escherichia coli* have been reported (21, 35, 46).

In a previous study, we found that mutants unable to synthesize UDP glucuronic acid (UDP-GlcUA), the nucleotide-sugar donor for the terminal sugar in the type 2 repeat unit, not only lacked GlcUA in the repeat unit but also produced low levels of polysaccharide that failed to be transferred to the cell wall. Each of these isolates also contained a secondary suppressor mutation (61). The suppressor mutations consistently mapped to genes within the capsule locus, primarily *cps2E*, and were indicative of lethality in the absence of reduced capsule synthesis. Suppressor mutations also occured with primary mutations that eliminated transport of repeat units by the Wzx flippase or their polymerization by the Wzy polymerase (61). In each case, the lethality was attributed to a buildup of lipid-linked repeat units which either destabilized the membrane or sequestered Und-P in the capsule pathway and away from essential pathways such as the peptidoglycan synthesis pathway.

Here, we use a genetic approach to identify the first committed, physiologically irreversible step in serotype 2 capsule synthesis. We also provide biochemical evidence for the functions of Cps2L, Cps2N, Cps2M, Cps2O, and Cps2T (also annotated WchF

[1,39]). The findings indicate that Cps2T is the β1-4 rhamnosyltransferase that adds the second sugar of the repeat unit, and this is the committed step in type 2 capsule synthesis.

MATERIALS AND METHODS

Materials. UDP-[^{14}C]Glc (293 mCi/mmol) was obtained from Amersham Biosciences. Inorganic pyrophosphatase, NAD_, NADH, dTTP, Glc-1-P, and dTDP-Glc were obtained from Sigma-Aldrich. Nonidet P-40 (NP-40) was obtained from Pierce Chemical Company. Solid-phase extraction (SPE) carbograph (150 mg/4 ml) columns were obtained from Grace Davison Discovery Sciences.

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 Dickinson] containing 3% defibrinated sheep blood [Colorado Serum Company]). Broth cultures of *S. pneumoniae* were grown statically at 37°C. Plate cultures were incubated at 37°C in a candle jar. *E. coli* strains (see Table S1 in the supplemental material) were grown at 37°C in L broth (1% tryptone, 0.5% yeast extract, 0.5% NaCl, 0.1% glucose) with shaking or on L-agar plates (15 g agar/liter L broth). The plasmids and primers used in this study are listed in Tables S1 and S2 in the supplemental material, respectively. Media were supplemented with the following antibiotics when 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).

TABLE 1. *S. pneumoniae* strains used in this study

Strain(s)	Properties[a]	References or source
AM1000	D39 Δ(*cps2A* - *cps2I*); Cps^-	(38)
BX515	D39 Δ*cps2K* repair; Cps^r ; retains *cps2E*$^{5924\ G\rightarrow T}$ (G303V)	(61)
D39	Type 2 parent strain; Cps^+	(3)
DJ901, DJ902, DJ910, DJ911, DJ973, DJ975, DJ977	pDJ078 x D39; Cps^- ; *Δcps2T* mutants isolated from five independent transformations; Km^R	This study
DJ904, DJ914, DJ916	pDJ086 x D39; Cps^- ; *Δcps2F* mutants with respective *cps2E* mutations: *cps2E* ^{5766insA} (E240* Frameshift); *cps2E*^{5924insG} (G293* Frameshift); *cps2E* $^{5718\ A\rightarrow G}$ (H258R); isolated from two independent transformations; Km^R	This study
DJ918, DJ919, DJ921	pDJ131 x D39; Cps^r ; *Δcps2G* mutants with respective *cps2E* mutations: *cps2E* $^{6059G\rightarrow C}$ (M338I); *cps2E* $^{5964C\rightarrow A}$ (R307S); *cps2E* $^{5529G\rightarrow T}$ (A162S); isolated from two independent transformations; Km^R	This study
DJ930, DJ943, DJ944	pDJ184 x D39; Cps^- ; *Δcps2I* mutants with respective *cps2E* mutations: *cps2E* $^{6267C\rightarrow T}$ (Q408*stop); *cps2E* $^{6222C\rightarrow A}$ (Q393K); *cps2E* ^{5221insCT} (M59*Frameshift); isolated from two independent transformations; Km^R	This study
DJ960	pDJ192 x DJ910; Cps^+ ; *Δcps2T* repair	This study
DJ965	pDJ192 x DJ901; Cps^+ ; *Δcps2T* repair	This study

[a] *cps2E* superscripts indicate locations of mutations compared to the sequence of GenBank accession no. AF026471. Amino acid changes are indicated in parentheses. Cpsr, reduced capsule levels; del, deletion; ins, insertion; →, nucleotide change; Km^r, kanamycin resistant; Ap^r, ampicillin resistant; Em^r, erythromycin resistant. Plasmids are listed in Table S1 in the supplemental material.

TABLE S1. Additional strains and plasmids used in this study

Strain(s) or plasmid(s)	Properties[a]	References or source
E. coli strains		
BL21-AI	F^- *ompT hsdSB*(r_B^- m_B^-) *gal dcm araB::*T7 RNAP-*tetA*	Invitrogen
DB11	*met thi gal hasdR nal rif*	(3)
DJ005	BL21-AI (pDJ005); full length Cps2I; Ap^R	This study
DJ009	BL21-AI (pDJ009); full length Cps2T; Ap^R	This study
DJ011	BL21-AI (pDJ011); full length Cps2G; Ap^R	This study
DJ014	BL21-AI (pDJ014); full length Cps2M with N-terminal-His_6; Ap^R	This study
DJ015	BL21-AI (pDJ015); full length Cps2N with N-terminal-His_6; Ap^R	This study
DJ016	BL21-AI (pDJ016); full length Cps2O with N-terminal-His_6; Ap^R	This study
DJ017	BL21-AI (pDJ017); full length Cps2L with N-terminal-His_6; Ap^R	This study
DJ052	BL21-AI (pDJ052); full length Cps2T with N-terminal-His_6; Ap^R	This study
DJ056	BL21-AI (pDJ056); full length Cps2G with N-terminal-His_6; Ap^R	This study
DJ078	DB11(pDJ078); *cps2T* deletion construct; Em^R	This study
DJ086	DB11(pDJ086); *cps2F* deletion construct; Em^R	This study
DJ089	BL21-AI (pDJ089); full length Cps2I with N-terminal-His_6; Ap^R	This study
DJ131	DB11(pDJ131); *cps2G* deletion construct; Em^R	This study
DJ184	DB11(pDJ184); *cps2I* deletion construct; Em^R	This study
DJ192	DB11(pDJ192); Δ*cps2T* repair construct; Em^R	This study
DJ204	BL21-AI (pDJ204); full length Cps2F, replaced GTG start with ATG; Ap^R	This study
DJ205	BL21-AI (pDJ205); full length Cps2F, replaced GTG start with ATG, N-terminal-His_6; Ap^R	This study
KJ4152	BL21-AI (pKJ4152); full length Cps2E; Ap^R	This study
RC124	BL21-AI (pET-20b); Ap^R	(1)
Plasmids		
pCR TOPO 2.1	PCR cloning vector; Ap^R Km^R	Invitrogen
pDL276	*aphA-3* containing vector; Km^R	(2)
pDJ005	pET-20b/*cps2I* amplified from D39 using primers I7/I8	This study
pDJ009	pET-20b/*cps2T* amplified from D39 using primers T13/T14	This study

TABLE S1. (Continued)

Strain(s) or plasmid(s)	Properties[a]	References or source
pDJ011	pET-20b/*cps2G* amplified from D39 using primers G10/G11	This study
pDJ014	pET-16b/*cps2M* amplified from D39 using primers M4/M5	This study
pDJ015	pET-16b/*cps2N* amplified from D39 using primers N4/N5	This study
pDJ016	pET-16b/*cps2O* amplified from D39 using primers O3/O4	This study
pDJ017	pET-16b/*cps2L* amplified from D39 using primers L6/L7	This study
pDJ052	pET-16b/*cps2T* amplified from D39 using primers T13/T14	This study
pDJ056	pET-16b/*cps2G* amplified from D39 using primers G10/G11	This study
pDJ078	pJY4164 containing *cps2E* (amplified from D39 using primers E54/E55), *aphA-3* (amplified from pSF151 using primers DJ-01/DJ-02), and *cps2F* (amplified from D39 using primer T12/F12). Used for *cps2T* deletion	This study
pDJ086	pJY4164 containing *cps2T* (amplified from D39 using primers T10/T11), *aphA-3* (amplified from pSF151 using primers DJ-01/DJ-02), and *cps2G* (amplified from D39 using primers F7/G8). Used for *cps2F* deletion	This study
pDJ089	pET-16b/*cps2I* amplified from D39 using primers I6/J8	This study
pDJ131	pJY4164 containing *cps2F* (amplified from D39 using primers F8/F9), *aphA-3* (amplified from pSF151 using primers DJ-01/DJ-02), and *cps2H* (amplified from D39 chromosomal DNA using primer G9/H7). Used for *cps2G* deletion.	This study
pDJ184	pJY4164 containing *cps2H* (amplified from D39 using primers H8/F9), *aphA-3* (amplified from pSF151 using primers DJ-01/DJ-02), and *cps2J* (amplified from D39 chromosomal DNA using primer I9/J9). Used for *cps2I* deletion.	This study
pDJ192	pJY4164/*cps2ETF* amplified from D39 using primers E8/F12. Used to repair Δ*cps2T* mutants.	This study
pDJ204	pET-20b/*cps2F* amplified from D39 using primers F10/F12. GTG start codon changed to ATG.	This study

TABLE S1. (Continued)

Strain(s) or plasmid(s)	Properties[a]	References or source
pDJ205	pET-16b/*cps2F* amplified from D39 using primers F10/F12. GTG start codon changed to ATG.	This study
pET16b	Protein expression vector, Ap^R	Novagen
pET20b	Protein expression vector, Ap^R	Novagen
pJY4164	*S. pneumoniae* suicide vector; Em^R	(5)
pKJ4152	pET-20b/*cps2E* amplified from D39 using primers E8/E11.	This study
pSF151	Streptococcal shuttle vector, derivative of pDL276	(4)

[a] Km^R, kanamycin resistant; Ap^R, ampicillin resistant; Em^R, erythromycin resistant.

TABLE S2. Primers used in this study

Primer[a]	Sequence[b]	Description[c]
E8 (+)	*catATG*AATGGAAAAACAGTAAAGTC	*cps2E*$^{5046–5068}$
E11 (-)	*ctcgag*CTACTTCGCTCCATCTCTC	*cps2E*$^{6414–6395}$
E54 (+)	gcgc*ggtacc*ATGAATGGAAAAACAGTAAAGTCTTC	*cps2E*$^{5046-5071}$
E55 (-)	gcgc*gcggccgc*CTACTTCGCTCCATCTCTCATAAATAC	*cps2E*$^{6413-6384}$
F7 (+)	gcgc*ctcgag*GTGTGGTGCAAGTAATCGGTG	*cps2F*$^{8529-8549}$
F8 (+)	gcgc*ggtacc*gTGAGTAACAAGCAAATTGCGATTATG	*cps2F*$^{7622-7648}$
F9 (-)	gcgc*gcggccgc*CTAAATAAACATTAACTCACCGATTAC TTGC	*cps2F*$^{8566-8536}$
F10 (+)	*cataTG*AGTAACAAGCAAATTGCGATT	*cps2F*$^{7623-7648}$
F12 (-)	gcgc*ggatcc*CTAAATAAACATTAACTCACCGATTACTT GC	*cps2F*$^{8566-8536}$
G8 (-)	gcgc*ggatcc*CTATTTACCGTTTTCAATATATACCCC	*cps2G*$^{9628-9602}$
G9 (+)	gcgc*ctcgag*GGGGTATATATTGAAAACGGTAAATAG	*cps2G*$^{9602-9628}$
G10 (+)	*catATG*AAAATTAATTTTATCCTTCCATTTAAG	*cps2G*$^{8582-8611}$
G11 (-)	*gaattc*CTATTTACCGTTTTCAATATATACCCC	*cps2G*$^{9628-9602}$
H7 (-)	gcgc*ggatcc*TTATTTTTCTTGCTTAGTCAATCTCATTC	*cps2H*$^{10832-10804}$

TABLE S2. (Continued)

Primer[a]	Sequence[b]	Description[c]
H8 (-)	gcgc*gcggccgc*TTATTTTTCTTGCTTAGTCAATCTCATTC	*cps2H*$^{10832-10804}$
H9 (+)	gcgc*ggtacc*ATGCTCTCTCTATATACAGGAAATGGTG	*cps2H*$^{9669-9696}$
I6 (+)	*ctcgag*ATGACAAAAAGTATCTTATATTTTTTATCTAC ATC	*cps2I*$^{10962-10985}$
I7 (+)	*catATG*ACAAAAAGTATCTTATATTTTTTATCTACATC	*cps2I*$^{10962-10996}$
I8 (-)	*gaattc*TCAATTTTCTAGTTCCTTATATAGTTGCATG	*cps2I*$^{12119-12089}$
I9 (+)	gcgc*ctcgag*GATAAAAATAATGCCATGCAGATCATG	*cps2I*$^{12066-12083}$
J8 (-)	*ggatcc*TCAATTTTCTAGTTCCTTATATAGTTGCATG	*cps2J*$^{12117-12088}$
J9 (-)	gcgc*ggatcc*TTATGTTAGAAACTTTTTTAATTCACCAAT AATAC	*cps2J*$^{13525-13491}$
L6 (+)	*catATG*AAAGGTATTATTCTTGCGGGTG	*cps2L*$^{15510-15534}$
L7 (-)	*gctaagc*CTAGACTTCTCCAATCAAACGGAGC	*cps2L*$^{16389-16359}$
M4 (+)	*catATG*ACAGATAATTTTTTCGGAAAAATAC	*cps2M*$^{16380-16407}$
M5 (-)	*ggatcc*TTACAAATCTTCTTTTTTCAAAGGTTTTAC	*cps2M* $^{16973-16944}$
N4 (+)	*catATG*ACTGAATACAAAAATATTATCGTGAC	*cps2N*$^{16985-17014}$
N5 (-)	*ggatcc*TTATACTGTAATAATCTCCTGAGTCTTAGC	*cps2N*$^{18035-18006}$
O3 (-)	*ctcgag*TTATCTCACTTCTTGTTTGTAAAATTCTTG	*cps2O*$^{18952-18923}$
O4 (+)	*catATG*ATTTTAATTACAGGGGCAAATG	*cps2O*$^{18101-18125}$
T10 (+)	gcgc*ggtacc*ATGAAGAAGTCAGTTTATATCATTGGTTC	*cps2T*$^{6445-6473}$
T11 (-)	gcgc*gcggccgc*TTACTCACTTTTTCCCCCTTCAAAC	*cps2T*$^{7629-7605}$
T12 (+)	gcgc*ctcgag*GAAGAGATAGTGGTGGATTATGAGGAAG	*cps2T*$^{7576-7603}$
T13 (+)	*catATG*AAGAAGTCAGTTTATATCATTGGTTC	*cps2T*$^{6445-6473}$
T14 (-)	*ctcgag*TTACTCACTTTTTCCCCCTTCAAAC	*cps2T*$^{7629-7605}$
DJ-01 (+)	gcgc*gcggccgc*GAGGAAGGAAATAATAA	*aphA-3*$^{1729-1746}$
DJ-02 (-)	gcgc*ctcgag*GTACTAAAACAATTCATCCA	*aphA-3*$^{2543-2523}$

[a] Forward and reverse primers are indicated by + and – , respectively.
[b] Lower case letters indicate restriction enzyme sites inserted for digestion. Restriction enzyme sites are italicized.
[c] Superscripts indicate the start and end positions of the primers in the type 2 D39 capsule locus. (GenBank accession no. AF026471) and pDL276 *aphA-3* (GenBank accession no. AF216803)

Generation of glycosyltransferase mutants. In-frame deletions of the glycosyltransferase-encoding genes were generated by allelic replacement with a nonpolarKmresistance (*aphA-3*) cassette. The entire upstream and downstream genes flanking the target gene were individually amplified from *S. pneumoniae* D39 chromosomal DNA by PCRs. The PCR products were individually ligated into pCR2.1-TOPO vectors as directed in the TOPO TA cloning kit manual (Invitrogen). The fidelity of each PCR product was confirmed by DNA sequencing (University of Alabama [UAB] Center for AIDS Research [CFAR] DNA sequencing core or the Genomics Core Facility of the Heflin Center for Genomic Science). The native ribosomal binding sites were included in the downstream flanking genes. The resulting flanking genes were ligated to either side of the Km resistance cassette and subcloned into the *S. pneumoniae* suicide vector pJY4164 in a single reaction. The ligation reaction mixture was transformed into chemically competent *E. coli* DB11 cells. Deletion constructs were generated using the following primer pairs: for the Δ*cps2T* construct, *cps2E* flank (E54/E55) and *cps2F* flank (T12/F12); for the Δ*cps2F* construct, *cps2T* flank (T10/T11) and *cps2G* flank (F7/G8); for the Δ*cps2G* construct, *cps2F* flank (F8/F9) and *cps2H* flank (G9/H7); and for the Δ*cps2I* construct, *cps2H* flank (H9/H8) and *cps2J* flank (I9/J9). *aphA-3* was amplified from the pneumococcal shuttle vector pSF151 using primers DJ-01 and DJ-02. Upstream flanks were digested with KpnI and NotI, downstream flanks were digested with XhoI and BamHI, *aphA-3* was digested with NotI and XhoI, and pJY4164 was digested with KpnI and BamHI.

For pneumococcal transformations, deletion constructs were transformed into competent D39 cells (27) and Km-resistant isolates were selected. Isolates were screened

for Em sensitivity to identify transformants that resulted from allelic replacement (Km^r, Em^s) rather than insertion of the suicide plasmid (Km^r, Em^r).

The glycosyltransferase mutations were repaired by allelic replacement using constructs that contained the relevant gene and full-length upstream and downstream genes. PCR products were ligated into pCR2.1-TOPO vectors, sequenced, subcloned into pJY4164, and transformed into *E. coli* DB11, as described above. Primer pairs and restriction enzymes for the Δ*cps2T* repair construct were E8/T14 and NdeI/XhoI.

For pneumococcal transformations, repair constructs were transformed into the relevant mutant background. Transformation mixtures were plated in the absence of selection. Isolates were screened for Km sensitivity, indicative of allelic replacement of the insertion.

Immunoblot and capsule analyses. For immunoblot analyses, proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS–12% PAGE) and transferred onto 0.22-_m-pore-size nitrocellulose membranes (GE Healthcare) using a semidry transfer apparatus (Fisher). Detection of Cps2E was performed as previously described (12). Other protein immunoblots were reacted with a tetra-His monoclonal antibody (MAb) (1:4,000 dilution; Qiagen), a MAb to the β subunit of *E. coli* RNA polymerase (1:2,000 dilution; Neoclone), or aMAb to the β subunit of ATP synthase (1:2,000 dilution; Abcam), as indicated. The secondary antibody was a goat anti-mouse immunoglobulin directly conjugated to alkaline phosphatase (1:5,000 dilution; Southern Biotech). Blots were washed and developed as previously described (4). *S. pneumoniae* fractionation and subsequent capsule immunoblot assays were performed as previously described (4), except that a 1:1,000 dilution of goat anti-rabbit

immunoglobulin directly conjugated to alkaline phosphatase was used instead of a streptavidin-biotin conjugate. Polyclonal antiserum to type 2 capsule (Statens Serum Institut) was used in the Quellung reaction (45) to visualize capsule microscopically. Methylpentose assays were performed as previously described (61).

S. pneumoniae fractionation and subsequent capsule immunoblot assays were performed as previously described (4), except that a 1:1,000 dilution of goat anti-rabbit immunoglobulin directly conjugated to alkaline phosphatase was used instead of a streptavidin-biotin conjugate. Polyclonal antiserum to type 2 capsule (Statens Serum Institut) was used in the Quellung reaction (45) to visualize capsule microscopically. Methylpentose assays were performed as previously described (61).

Purification and characterization of *S. pneumoniae* dTDP-Rha biosynthesis enzymes. To characterize dTDP-Rha biosynthesis enzymes, the sequences for *cps2L*, *cps2N*, *cps2M*, and *cps2O* were individually amplified from *S. pneumoniae* D39 chromosomal DNA by PCR using the following primer pairs: for *cps2L*, L6/L7; for *cps2N*, N4/N5; for *cps2M*, M4/M5; and for *cps2O*, O3/O4. The PCR products were ligated into pCR2.1-TOPO vectors and sequenced as described above. The respective sequences were subcloned into pET-16b plasmids (facilitating generation of N-terminal His tags) and transformed into chemically competent *E. coli* BL21-AI cells. The resulting strains, DJ017 (*cps2L*), DJ014 (*cps2M*), DJ015 (*cps2N*), and DJ016 (*cps2O*), were used for expression of the respective proteins. The pET plasmids used in this study contain a T7 RNA polymerase promoter. *E. coli* BL21-AI contains a chromosomally encoded T7 RNA polymerase under the control of an arabinose-inducible promoter.

For protein expression, 250-ml cultures were grown under the appropriate antibiotic selection to mid-exponential phase (cell density, approximately 4.5_108 CFU/ml) and then induced for 4 h with 0.2% arabinose. Cultures were centrifuged at 3,000 x *g* for 20 min, and the pellet was suspended in 5 ml of wash buffer (300 mM NaCl, 50 mM NaH_2PO_4, pH 7.0). Cells were sonicated on ice, using a Sonic Dismembrator 300 apparatus (Fischer) at 50% power, 5 times for 10 s each time with 30-s pause intervals. The lysate was then centrifuged at 8,000 x *g* for 10 min to pellet cellular debris. The cell-free lysate was passed over a Talon cobalt-metal affinity resin (Clontech) to purify the respective proteins. Native purification procedures were followed as directed in the manufacturer's specifications. Eluates containing purified His-tagged proteins were concentrated using Amicon Ultra 10-kDa centrifugal filters (Millipore). The retained proteins were washed twice with 500 μl of 50 mM Tris-acetate, pH 7.6, and then suspended in 100 μl of 30% glycerol in 50 mM Trisacetate, pH 7.6. Protein concentrations were calculated from absorbances at 280 nm using Beer's law. Protein extinction coefficients were calculated using the Scripps Research Institute Protein Calculator Program (http: //www.scripps.edu/cdputnam/protcalc.html). Enzymes were stored at -80°C and were active for several months.

To assay Glc-1-P thymidylyltransferase (Cps2L) activity, 300-_l reaction mixtures containing 30 mM HEPES, pH 7.6, 3 mM dTTP, 3 mM Glc-1-P, 9mM $MgCl_2$, and 0.3 μMCps2L were incubated for 1 h at 37°C. Proteins were removed using Amicon Ultra 10-kDa centrifugal filters. The filtrate was analyzed by high-performance liquid chromatography (HPLC) with a CarboPac PA-100 or PA-1 column utilizing UV detection at 260 nm and a gradient of ammonium acetate (0 to 1 M) as an eluent.

To assay dTDP–D-Glc-4,6-dehydratase (Cps2N) activity, 300 µl reaction mixtures containing 30mM HEPES, pH 7.6, 3mM dTDP-Glc, 1mM NAD^+, and 0.3 µM Cps2N were incubated for 1 h at 37°C. Proteins were removed, and HPLC analysis was performed as described above.

To further characterize Cps2L and Cps2N activities, spectrophotometric analyses were used to monitor NADH accumulation as a by-product of dTDP–6-deoxy-α-D-*xylo*-4-hexulose synthesis. Synthesis of dTDP–6-deoxy-α-D-*xylo*-4-hexulose using dTDP–D-Glc-4,6-dehydratase yields a characteristic maximum absorbance at 320 nm in an alkaline solution (46). Products were generated in a 500-µl reaction mixture containing 30 mM HEPES, pH 7.6, 3 mM dTTP, 3 mM Glc-1-P, and 9 mM MgCl2 with 0.3 µMCps2L and/or 0.3 µMCps2N. The reaction mixtures were incubated at 37°C. At 30-min intervals, 25-µl aliquots were removed and placed in 775 µl of 100 mM NaOH. Incubations were continued for an additional 15 min before determining the absorbance at 320 nm.

dTDP–6-deoxy-D-*xylo*-4-hexulose-3,5-epimerase (Cps2M) and dTDP-6-deoxy-L-*lyxo*-4-hexulose-reductase (Cps2O) activities were analyzed in a combined spectrophotometric assay measuring the oxidation of NADH at 340 nm. Kinetic experiments were carried out with dTDP–6-deoxy-α- D-*xylo*-4-hexulose (filtrate from the above-mentioned Cps2N reaction), 1 mM NADH, 0.3 µM Cps2M, and/or 0.3 µM Cps2O. The reactions were carried out in 300-µl volumes, and the reaction mixtures were incubated for 1 h at 37°C.

Synthesis, purification, and quantification of dTDP-Rha. dTDP-Rha was synthesized in two steps. First, dTDP-Glc was synthesized in a 350-µl reaction mixture containing 50 mM HEPES, pH 7.6, 30 mM $MgCl_2$, 3.4 mM dTTP, 3.4 mM Glc-1-P, 2

units of inorganic pyrophosphatase, and 1 μM Cps2L. After an incubation of 1 h at 37°C, the proteins were removed by filtration as described above. The filtrate was used in the second step and supplemented with 0.1mM NAD^+, 6mM NADH, 1 μM Cps2N, 1 μM Cps2M, and 1 μM Cps2O in a final volume of 350 μl. The reaction mixture was incubated for 1 h at 37°C, the proteins were filtered, and dTDP-Rha was separated by HPLC as described above. Fractions containing dTDP-Rha were pooled and then desalted on an Extract-Clean SPE carbograph (150 mg/4 ml) column. The column was first equilibrated by adding 4ml of 80% acetonitrile–0.1% trifluoroacetic acid, followed by 10 ml of water. The sample was then added to the column, followed by another 10 ml of water. dTDP-Rha was eluted with 1.5 ml of 25% acetonitrile–0.1% trifluoroacetic acid. The effluent was freeze-dried using a SpeedVac apparatus (Savant). The dried sample was suspended in 0.5 ml anhydrous methanol and again freeze-dried. This process was repeated 5 times, and the dried sample was stored at -20°C. For use in glycosyltransferase assays, samples were resuspended in 100 mM Tris-acetate, pH 7.6, and dTDP-Rha concentrations were calculated from the absorbances at 260 nm using Beer's law and the thymidine extinction coefficient (7.4×10^3 M^{-1} cm^{-1}). For electrospray ionization tandem mass spectrometry (ESI-MS/MS) analysis (see below), samples were suspended in the mobile phase prior to analysis.

GC/MS and ESI-MS/MS analyses. The identity of dTDP-Rha was confirmed using gas chromatography-mass spectrometry (GC-MS) and ESI-MS/MS. A 1-μl aliquot of a 10 mMsolution of Glc and Rha standards, along with 1 μl of purified dTDP-Rha (approximately 1 mM), was dried in a SpeedVac apparatus for sugar component analyses. Dried samples were subjected to methanolysis (500μl of 3N methanolic HCl at 80°C) for

16 h, dried, and washed 5 times with 500 μl of anhydrous methanol. The dried samples were trimethylsilylated by reacting them with 100 μl Tri-Sil reagent (Pierce) for 30 min at 80°C. The reaction products were analyzed on a GC-MS (Varian 4000; Agilent Technologies, Santa Clara, CA) fitted with a 30-m (inner diameter, 0.25 mm) VF-5ms capillary column. The column temperature was maintained at 160°C for 3 min and then increased to 260°C at 3°C/min and finally held at 260°C for 2 min. The effluent was analyzed by MS using the electron impact ionization mode.

For ESI-MS/MS, 1 μl of purified dTDP-Rha (approximately 1 mM) was injected by direct infusion into an API4000 (Applied Biosystems) triple-quadrupole mass spectrometer operated under the negative electrospray ionization mode. The ion spray voltage was set to -4,000 V. The flow injection mobile phase was 50% acetonitrile–0.1% formic acid, and the flow rate was 20 μl/min. The collision energy, collision gas, declustering potential, and nebulizing gas parameters were set to -40, 4, -40, and 15, respectively.

Glycosyltransferase expression and isolation. Sequences for *cps2E*, *cps2T*, *cps2F*, *cps2G*, and *cps2I* were individually amplified from *S. pneumoniae* D39 chromosomal DNA by PCR using the following primer pairs: for *cps2E*, E8/E11; for *cps2T*, T13/T14; for *cps2F*, F10/F12; for *cps2G*, G10/ G11; and for *cps2I*, I7/I8 (or J8/I6). The GTG start of *cps2F* was changed to ATG to facilitate proper *in vitro* expression. Cloning with primer pair I7/I8 requires a partial digest with NdeI to acquire full-length Cps2I. PCR products ligated into pCR2.1-TOPO were sequenced for verification. Appropriate sequences were subcloned into pET-16b or pET-20b and transformed into chemically competent *E. coli* BL21-AI cells as described above. The

resulting strains were used for expression of the respective proteins. Strains KJ4152 (*cps2E*), DJ009 (*cps2T*), DJ204 (*cps2F*), DJ011 (*cps2G*), and DJ005 (*cps2I*) containing the non-His-tagged proteins were used in glycosyltransferase assays, while DJ052 (*cps2T*), DJ205 (*cps2F*), DJ056 (*cps2G*), and DJ089 (*cps2I*) containing the His-tagged proteins were used in protein localization experiments. Cps2T, Cps2F, Cps2G, and Cps2I were expressed as described for the dTDP-Rha biosynthetic enzymes. Following centrifugation of the cultures, cell pellets were frozen and subsequently used to isolate membrane and cytoplasmic fractions. Cultures expressing Cps2E were processed as previously described (12). For all, frozen cell pellets were thawed at room temperature and suspended in 6 ml of spheroplast buffer (10 mM Tris- HCl [pH 7.5], 20% sucrose, 10 mM EDTA) containing 0.4 mg/ml lysozyme, 0.5 μg/ml pepstatin, 0.7 μg/ml leupeptin, 1 mM phenylmethylsulfonyl fluoride (PMSF), and 1 mM dithiothreitol (DTT). The cellular suspension was incubated at 4°C with constant mixing for 4 h. Spheroplasts were sedimented by centrifugation at 8,000 x *g* for 10 min and lysed by suspension in 6 ml of sterile water containing 10 mM EDTA, 1 mM PMSF, and 1 mM DTT. Suspensions were sonicated on ice as described above. Cellular debris was sedimented by centrifugation at 8,000 x *g* for 20 min. The supernatant, containing both the cytoplasmic and membrane fractions, was centrifuged at 50,000 x *g* for 20 min. The supernatant was saved as the cytoplasmic fraction and centrifuged once more at 50,000 x *g* for 20 min to reduce residual membranes. The pellet (membranes) was washed two times with 3 ml of 100 mM Tris-acetate (pH 7.5) containing 10% glycerol and 1 mM DTT. The final membrane pellet was suspended in 500 μl of 100mM Tris-acetate (pH 7.5) containing 10% glycerol

and 1 mM DTT. Total cytoplasmic and membrane proteins were quantified using the Bradford Bio-Rad protein dye assay.

The interactions of the glycosyltransferases with the membrane were tested by suspending pelleted membranes in 30 μl of 100 mM Na2CO3 (pH 11.5). The mixture was centrifuged at 50,000 x *g* for 20 min, and the entire supernatant was saved. This process was repeated twice; after the final wash was collected, the remaining membranes were suspended in 30 μl of 100 mM Na_2CO_3.

Glycosyltransferase assays. Glycosyltransferase assays were modified from those described elsewhere (12, 34). Briefly, membranes containing 3 μg of total protein were incubated for 1 h at 10°C in reaction mixtures containing 10 mM MnCl2, 1 mM DTT, 0.025 μCi UDP-[^{14}C]Glc (293 mCi/mmol), and, where indicated, 0.1 mM dTDP-Rha. The final volume was brought to 75 μl with 100 mM Tris-acetate, pH 7.6. For reaction mixtures containing independent membrane preparations, membranes were preincubated with 0.008% NP-40 (Pierce) for 10 min, prior to the addition of metal ions and nucleotide-sugar donors. Reactions were stopped, and the organic phase was extracted as described previously (12). Glycolipids were either preserved for later use, as described for Und-P-P-[^{14}C]Glc (see below), or subjected to mild-acid hydrolysis to liberate the saccharide moiety for separation by thin-layer chromatography (TLC), as has been used previously to analyze glycosyltransferase activity (12, 34, 47). Hydrolysis was performed by suspending the dried organic phase in 80 μl of 20 mM HCl and incubating at 70°C for 20 min. Hydrolyzed samples were dried, suspended in 10 μl of chloroform-methanol (1:1), and applied as spots on silica-coated 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).

Synthesis of Und-P-P-[14C]Glc and assay of Cps2T activity in the absence of Cps2E. Und-P-P-[^{14}C]Glc was synthesized as previously described (12) with a few modifications. Briefly, membranes (10 μg total protein) containing Cps2E were incubated as described above for the glycosyltransferase assays. Extracted glycolipids were dried and suspended in 100 μl of water (or 100 mM Tris-acetate, pH 7.6) containing 0.1% NP-40. This solution was used as the source of Und-P-P-[^{14}C]Glc in subsequent glycosyltransferase reactions. Membranes (3 μg total protein) from DJ009 containing non-His-tagged Cps2T or from RC124, a vector only *E. coli* strain, were incubated in reaction mixtures containing 0.008% NP-40, 10mM MnCl2, 1 mM DTT, 25 μl of the Und-P-P-[^{14}C]Glc solution, and 0.1 mM dTDP-Rha. Reaction products were visualized as described above.

RESULTS

Genetic evidence that Cps2T performs the committed step in serotype 2 capsule synthesis. From our previous study (61), we knew that mutations resulting in loss of either the terminal sugar (a GlcUA branched from the backbone), Wzx-mediated transport, or Wzy-mediated assembly of the repeat unit could be obtained only in the presence of suppressor mutations that reduced repeat unit synthesis. The primary mutations were postulated to be lethal because they resulted in the accumulation of lipid-linked repeat units that could not be readily disassembled; i.e., capsule synthesis in these

mutants had proceeded beyond a committed step. Since suppressor mutations were required to obtain the Wzx mutants, the committed step must occur prior to transport of the repeat unit across the membrane and thus be performed by a glycosyltransferase. To identify the glycosyltransferase that performs the committed step, we took advantage of the requirement for secondary suppressor mutations to prevent the lethal accumulation of lipid-linked repeat units. We hypothesized that mutations affecting enzymes downstream of the committed step would be obtainable only in the presence of suppressor mutations, whereas mutation of the enzyme performing the committed step should not require suppressor mutations.

We first constructed deletions of *cps2T, cps2F, cps2G,* and *cps2I* in *S. pneumoniae* by replacing the respective open reading frames with a nonpolar Km resistance cassette. For each gene, multiple mutants were derived from at least two independent transformation reactions with D39. All mutants displayed the small, rough colony morphology indicative of loss of capsule, and no capsule was detected on any of the mutants using type 2-specific antiserum with intact cells in the Quellung reaction (data not shown). During growth in liquid medium, the mutant cultures exhibited an abrupt decrease in optical density, as did other capsule-negative mutants, compared to that for the parent strain (shown in Fig. 2 for Δ*cps2T* and Δ*cps2G*). These mutants aggregated and precipitated out of the culture medium, consistent with contact of exposed hydrophobic cell surfaces resulting from the absence of a parental capsule.

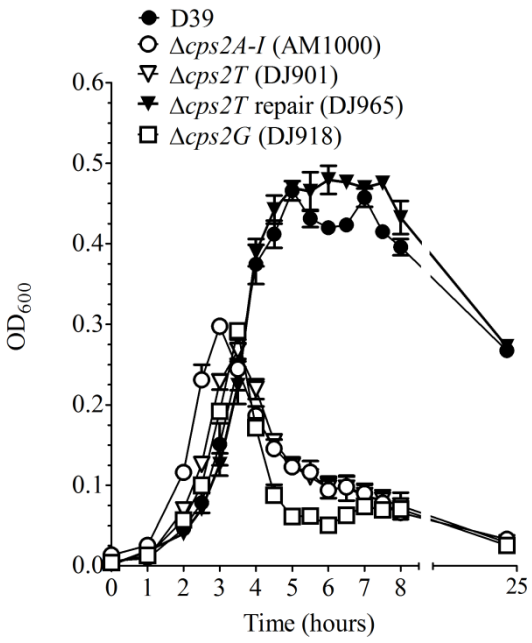

FIG 2. Growth of D39 and derivatives. Cultures grown in THY to an OD_{600} of 0.15 were diluted 1/10 and grown in THY. A_{600} readings were taken at the indicated time points. Data are the means and standard errors from two cultures. AM1000 is a capsule-negative mutant of D39.

To further investigate whether any capsule was detectable in specific cellular factions, the protoplast (membrane-containing) and cell wall fractions were individually extracted, separated by SDS-PAGE, transferred to a nitrocellulose membrane, and reacted with type 2-specific antiserum. Here, no capsule was detected in independent Δ*cps2T*, Δ*cps2F*, or Δ*cps2I* mutants (shown in Fig. 3 for a Δ*cps2T* mutant). For the Δ*cps2G* mutants, a very small amount of capsule was detected in the protoplast fraction, but none was detected in the cell wall fraction, even when the starting material was concentrated an additional 5-fold (Fig. 3 and data not shown).

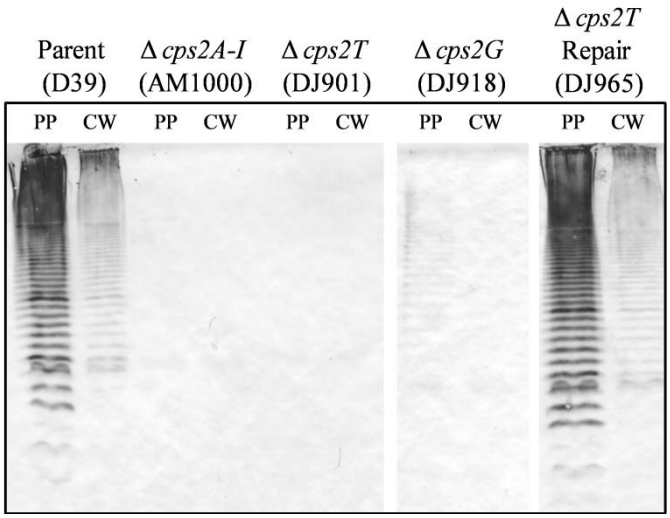

FIG 3. Capsule production by D39 and derivatives. Protoplast (membrane-containing) and cell wall fractions were extracted as described in the Materials and Methods. Samples were separated by SDS-12% PAGE and transferred to a nitrocellulose membrane. Capsule was detected using polyclonal antiserum against the type 2 capsule. Protoplasts (PP) were normalized to 30 μg of total protein and cell wall (CW) fractions were relative to their protoplast counterparts. AM1000 is a capsule-negative mutant of D39.

The reduction of CPS in the Δ*cps2G* mutants was further demonstrated by analysis of the methylpentose (rhamnose) levels of polymers extracted from whole cells. The methylpentose levels in strains D39 (parent, Cps^+), DJ918 (Δ*cps2G*), and AM1000 (Cps^-) were 1.4 ± 0.024, 0.13 ± 0.017, and 0.11±0.013 μg/10^8 CFU, respectively. The values represent the means ± standard errors from two assays performed in duplicate. The level for D39 was significantly different from the levels for *cps2G* mutant DJ918 and nonencapsulated strain AM1000 ($P < 0.001$), which were not different from each other, as determined using one-way analysis of variance with Bonferroni posttests. These results indicate that the immunoblot detection of capsule in *cps2G* mutants represents reduced CPS levels and not altered antibody reactivity from loss of an epitope. The ability of

Δ*cps2G* mutants to synthesize polymers indicated that Cps2G does not add a backbone sugar.

In our previous study, most of the suppressor mutations mapped to *cps2E* (61). We therefore sequenced this gene from multiple isolates of the glycosyltransferase mutants. All of the Δ*cps2F*, Δ*cps2G*, and Δ*cps2I* mutants (three from two independent reactions for each mutant) contained mutations in *cps2E* (Table 1). In each case, the *cps2E* mutations were different. Sequencing of *cps2E* from the Δ*cps2T* mutants (seven obtained from five independent reactions) revealed no mutations. Repair of independent Δ*cps2T* mutants resulted in restoration of the capsule, level of growth, and colony morphology to those for the parent (Fig. 2 and 3 and data not shown), demonstrating that these mutants did not contain secondary mutations that affected capsule synthesis. In contrast, repair of other deletion mutations did not restore parental phenotypes (data not shown). These results indicated that the phenotypes observed in the latter mutants were not solely due to the primary deletion mutations but were affected by the *cps2E* mutations and that suppressor mutations were required to obtain mutants lacking Cps2F, Cps2G, and Cps2I but not Cps2T, thus identifying the Cps2T reaction as the committed step.

During construction of the glycosyltransferase mutants, the Δ*cps2G* and Δ*cps2I* mutants were obtained at very low frequencies compared to those for the Δ*cps2T* and Δ*cps2F* mutants (Table 2). The frequencies for the former were increased when a recipient that already contained a *cps2E* point mutation that reduced Cps2E activity was used (Table 2), underscoring the need for a suppressor mutation to obtain *cps2G* and *cps2I* mutants. The fact that Δ*cps2F* mutants were obtained at relatively high frequencies, which were not increased further by a preexisting *cps2E* mutation, suggests differences in the detrimental effects of the *cps2F* mutation relative to those of *cps2G* and *cps2I*. These observations are considered in more detail in the Discussion.

Table 2 Frequencies of acquiring deletion mutations

	Transformant Frequency[a]	
Mutation	D39	BX515
Δ*cps2T* [b]	1.2×10^{-2} (9.2×10^{-3})	1.2×10^{-2} (4.2×10^{-3})
Δ*cps2F* [b]	4.2×10^{-3} (1.1×10^{-3})	2.8×10^{-3} (8.4×10^{-5})
Δ*cps2G* [b]	3.2×10^{-5} (1.1×10^{-6})	2.6×10^{-3} (5.4×10^{-4})[c]
Δ*cps2I* [b]	2.9×10^{-5} (1.7×10^{-5})	3.1×10^{-4} (3.2×10^{-5})[c]

[a] Recipients were the D39 parent and the D39 derivative BX515, which contains a *cps2E* point mutation

[b] Frequencies are the number of transformants per recipient, i.e. antibiotic resistant CFU/total CFU, expressed as the means (± standard errors) from two independent transformations. The complete data sets are provided as supplemental information in Table S3.

[c] Donor DNAs used for deletions were obtained from *E. coli* constructs indicated in Table S1.

[d] Significantly different ($P<0.05$) from D39, as determined using a two-tailed unpaired Student *t*-test.

TABLE S3. Frequencies of acquiring deletion mutations (Complete data set) Recipients were the D39 parent and the D39 derivative BX515, which contains a *cps2E* point mutation

		Transformant Frequency[a]	
Mutation	Exp #[b]	D39	BX515
Δ*cps2T* [c]	1	290/116,000 (2.5×10^{-3})	410/25,000 (1.6×10^{-2})
	2	710/34,000 (2.1×10^{-2})	380/48,000 (7.9×10^{-3})
	mean	1.2×10^{-2} (9.2×10^{-3})	1.2×10^{-2} (4.2×10^{-3})
Δ*cps2F* [c]	1	280/92,000 (3.0×10^{-3})	74/27,000 (2.7×10^{-3})
	2	270/51,000 (5.3×10^{-3})	64/22,000 (2.9×10^{-3})
	mean	4.2×10^{-3} (1.1×10^{-3})	2.8×10^{-3} (8.4×10^{-5})
Δ*cps2G* [c]	1	3/90,000 (3.3×10^{-5})	54/26,000 (2.1×10^{-3})
	2	2/64,000 (3.1×10^{-5})	79/25,000 (3.2×10^{-3})
	mean	3.2×10^{-5} (1.0×10^{-6})	2.6×10^{-3} (5.4×10^{-4})[d]
Δ*cps2I* [c]	1	4/87,000 (4.6×10^{-5})	7/25,000 (2.8×10^{-4})
	2	1/88,000 (1.1×10^{-5})	11/32,000 (3.4×10^{-4})
	mean	2.9×10^{-5} (1.7×10^{-5})	3.1×10^{-4} (3.2×10^{-5})[d]

[a] within each mutation group, data above the line are the number of antibiotic resistant CFU/total CFU for two independent experiments, with the calculated frequencies shown in parentheses. The data below the lines are the means of the frequencies from the two experiments, with ± standard errors shown in parentheses.
[b] experiment number for two independent experiments using different batches of competent *S. pneumoniae* and different preparations of donor DNA.
[c] donor DNAs used for deletions were obtained from *E. coli* constructs indicated in Table S1.
[d] significantly different ($P<0.05$) from D39, as determined using a two-tailed unpaired Student *t*-test.

Synthesis of dTDP-rhamnose and biochemical evidence for Cps2L, Cps2N, Cps2M, and Cps2O functions. To biochemically demonstrate activities for the dTDP-Rha biosynthetic enzymes (Fig. 4A) and to generate a source of dTDP-Rha for use in *in vitro* glycosyltransferase assays, *cps2L*, *cps2M*, *cps2N*, and *cps2O* were individually cloned from *S. pneumoniae* D39 and expressed with N-terminal His tags in *E. coli* (Fig.

4B). Enzymatic activities of the purified proteins were initially demonstrated using indirect assays that followed reduction of NAD^+ to NADH in reaction mixtures containing Cps2L and Cps2N (see Fig. S1A in the supplemental material) or oxidation of NADH in reaction mixtures containing Cps2M and Cps2O (see Fig. S1B in the supplemental material). The products generated from individual reactions were analyzed using HPLC (Fig. 4C). The Cps2L product was confirmed to be dTDP-Glc by its retention time, which was similar to that of the standard dTDP-Glc. The Cps2N and Cps2M reaction products had a delayed retention time compared to that of their precursors, similar to other dTDP-keto intermediates analyzed with these methods (16). The shorter retention time for the Cps2O product relative to dTDP-Glc is consistent with that reported for dTDP-Rha in other publications (23, 50). Sugar component analysis of the purified HPLC peak corresponding to the Cps2O reaction product confirmed the presence of Rha and the absence of Glc (Fig. 4D), while the mass of the purified Cps2O reaction product was confirmed through ESI-MS/MS to be that of dTDP-Rha (see Fig. S2 in the supplemental material).

***cps2T* encodes the β1-4 rhamnosyltransferase that adds the second sugar to the repeat unit.** We previously demonstrated that Cps2E initiates serotype 2 capsule repeat unit synthesis by transferring Glc-1-P to Und-P, generating Und-P-P-Glc (12). With the premise of a stepwise addition for each sugar residue, the next step is the addition of a single β1-4 Rha residue (Fig. 1). To determine which of the predicted glycosyltransferases accomplishes this step, we first cloned and expressed Cps2T, Cps2F, Cps2G, and Cps2I in *E. coli*. We then performed cellular fractionation experiments and found that Cps2T, Cps2F, Cps2G, and Cps2I were contained in the membrane fraction

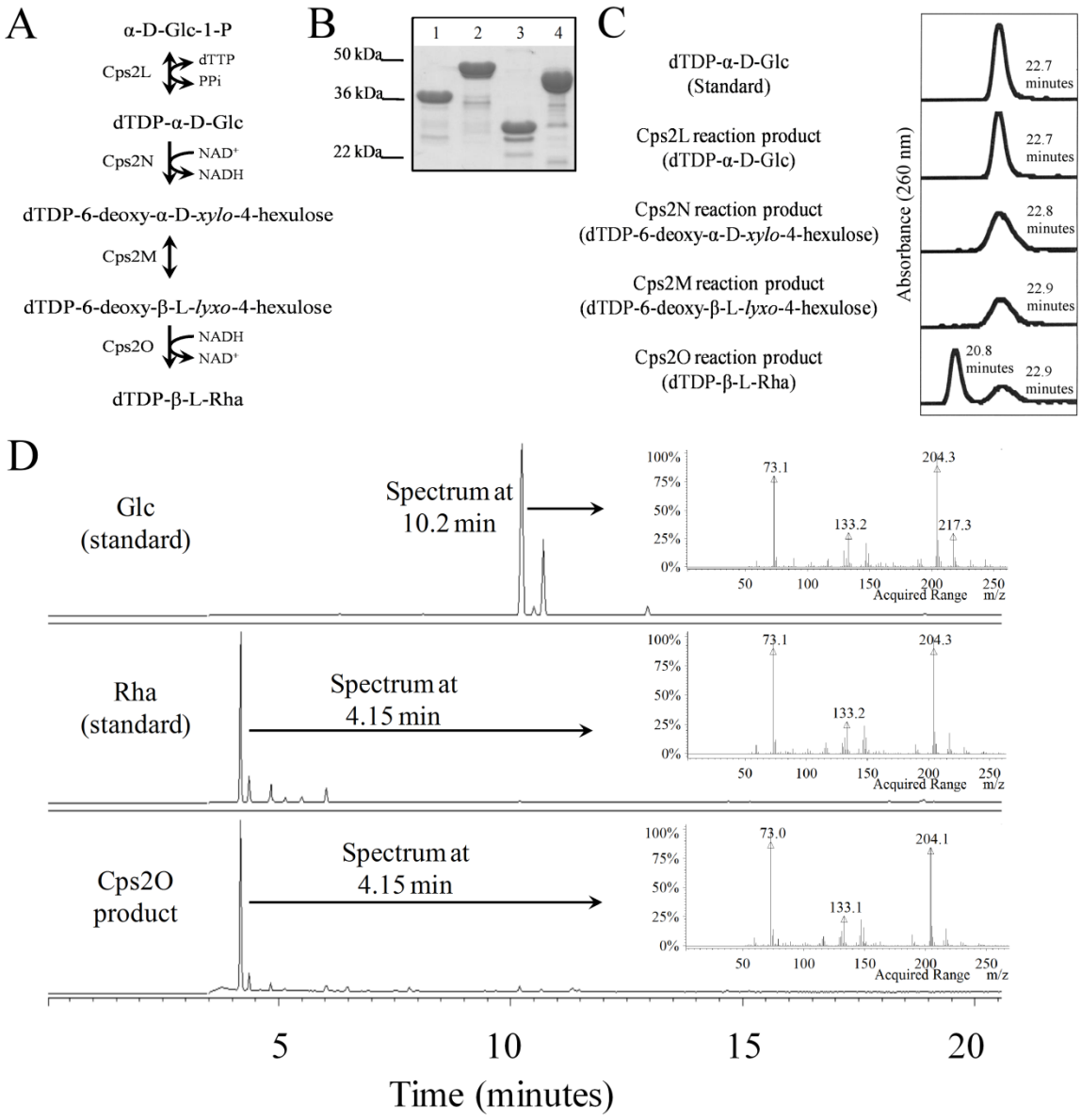

FIG 4. dTDP-Rha biosynthesis. (A) dTDP-Rha biosynthesis pathway in *S. pneumoniae* type 2. (B) dTDP-Rha biosynthesis enzymes were expressed in *E. coli* and purified using an incorporated N-terminal His tag. Purified proteins (0.1 μg) were separated by SDS–12% PAGE and stained with Coomassie brilliant blue. Lanes: 1, Cps2L (calculated molecular mass, 34.7 kDa); 2, Cps2N (41.5 kDa); 3, Cps2M (24.8 kDa); 4, Cps2O (34.8 kDa). Apparent molecular masses from standards are presented on the left. (C)HPLCanalyses of the Cps2L, Cps2N,Cps2M,andCps2Oreaction products. Samples analyzed, listed from the top to bottom, are dTDP-Glc standard, Cps2L, Cps2N, Cps2M, and Cps2O reaction products. (D) GC-MS sugar component analyses of Glc standard, Rha standard, and purified Cps2O product.

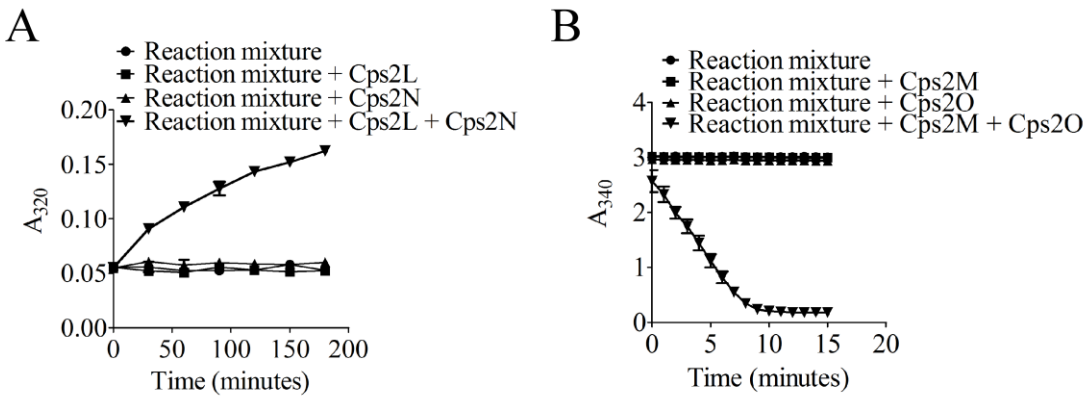

FIG S1. Kinetic assays for Cps2N and Cps2O. Assays were performed as described in the Materials and Methods. The reaction mixture for Cps2N activity (A) contained dTTP, Glc-1-P, NAD^+, and the indicated enzyme(s). Activity was measured as the generation of NADH. The reaction mixture for Cps2O activity (B) contained the Cps2N product, NADH, and the indicated enzyme(s). Activity was measured as the oxidation of NADH. Data points represent means ± standard errors from three reactions.

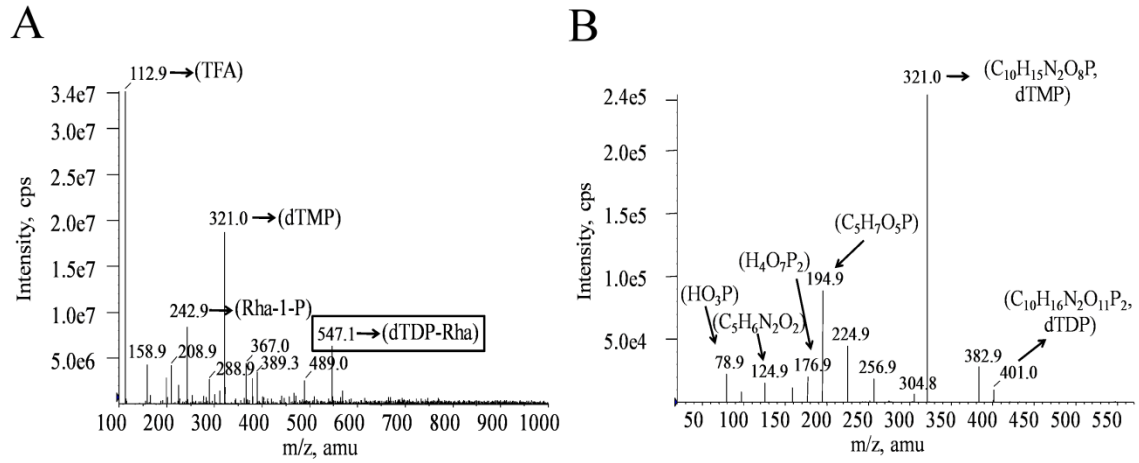

FIG S2. ESI-MS/MS of HPLC purified Cps2O product. One µl of purified dTDP-Rha (approximately 1 mM) was injected by direct infusion into a triple quadruple mass spectrometer operated under the negative electrospray ionization mode. (A) Precursor ion spectra of the purified Cps2O product (ESI-MS). Ion 547.1 corresponds to dTDP-Rha . (B) The product ion spectra (ESI-MS/MS) of ion 547.1. Product ions indicated are consistent with calculated fragments of dTDP-Rha. Ion intensity is presented as counts per second (cps).

(Fig. 5A and B). In contrast to Cps2E, a known integral membrane protein, large amounts of Cps2T, Cps2F, Cps2G, and, to a lesser extent, Cps2I dissociated from the membranes by washing with Na_2CO_3 (pH 11.5), a procedure that releases peripherally associated membrane proteins (19) (see Fig. S3 in the supplemental material).

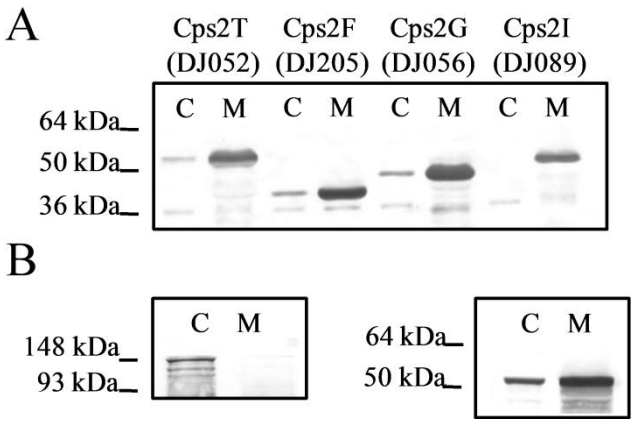

FIG 5. Cps2T, Cps2F, Cps2G, and Cps2I localization in *E. coli*. (A) N-terminal His-tagged glycosyltransferases were expressed in *E. coli*. The cytoplasmic and membrane fractions were extracted and normalized as described in Materials and Methods. Three micrograms of total protein for each fraction was separated by SDS-12% PAGE, transferred to a nitrocellulose membrane, and probed with anti-tetra-His antibody. The calculated molecular masses for Histagged Cps2T, Cps2F, Cps2G, and Cps2I are 47.6, 38.7, 42.9, and 46.8 kDa, respectively. (B) Fractionation controls, using representative preparations from DJ052 (Cps2T), were probed with MAb to the β subunit of *E. coli* RNA polymerase (150 kDa) for cytoplasmic protein detection (left) and with MAb to the β subunit of ATP synthase (50 kDa) for membrane protein detection (right). C, cytoplasmic fraction; M, membrane fraction.

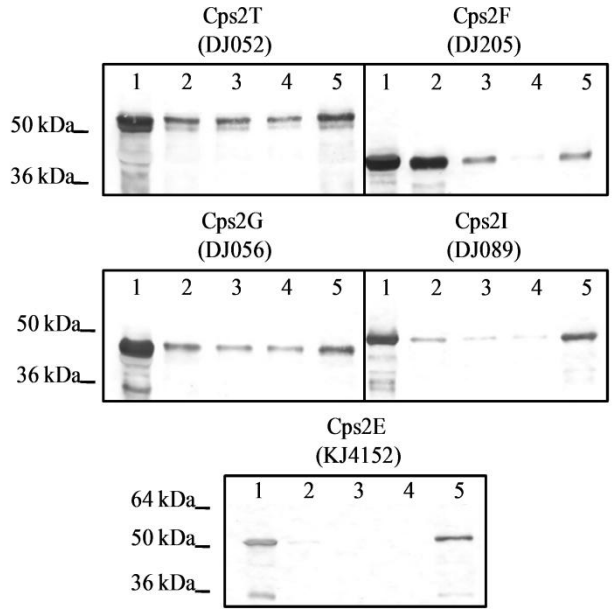

FIG S3. Cps2E, Cps2T, Cps2F, Cps2G, and Cps2I membrane association in *E. coli*. Membrane fractions (3 μg of total protein) were sequentially washed with 30 μl of 100 mM Na_2CO_3 (pH 11.5). Lanes: 1, 3 μg of total membrane protein from the indicated strains; 2, first wash with 100 mM Na_2CO_3 (the entire 30 μl was loaded into the gel); 3 and 4, second and third washes with 100 mM Na_2CO_3, respectively; 5, remaining proteins associated with the membrane. Cps2T, Cps2F, Cps2G, and Cps2I blots were probed with α-Tetra-His; the Cps2E blot was probed with α-Cps2E.

To assess glycosyltransferase activities, we used Cps2E-containing membranes in reaction mixtures with membranes containing the respective glycosyltransferases and nucleotide sugars. Under these conditions, we found it essential to add a small amount of nonionic detergent (0.008% NP-40) for enzyme activity. Glycolipids formed in the reactions were extracted and subjected to mild-acid hydrolysis to liberate the saccharide moiety, and the products were separated by TLC. In the presence of UDP-[^{14}C]Glc only, a single product consistent with the incorporation of Glc was observed for all reactions (Fig. 6A).

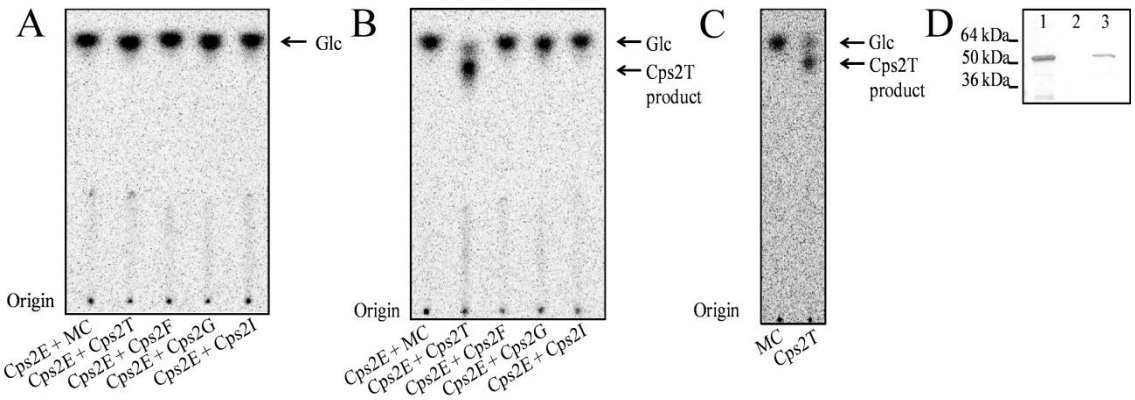

FIG 6. Cps2T rhamnosyltransferase activity. Cps2E-containing membranes (3 μg total membrane protein) were incubated with membranes (3 μg total protein) containing the indicated non-His-tagged glycosyltransferases or membrane controls (MC; 3 μg total protein from vector-only *E. coli* strain RC124) and UDP-[^{14}C]Glc only (A) or UDP-[^{14}C]Glc and dTDP-Rha (B). Reaction mixtures (total volume, 75 μl) were incubated for 1 h at 10°C and contained 0.008% NP-40, 10mM $MnCl_2$, 1mM DTT, 0.025 μCi UDP[^{14}C]Glc (293 mCi/mmol), and, for panels B and C, 0.1mM dTDP-Rha. Glycolipids were extracted, mild-acid hydrolyzed to liberate the saccharide moiety, and separated by TLC as described in Materials and Methods. The migration of the product designated Glc was equivalent to that of a [^{14}C]Glc standard. (C and D) Cps2T activity in the absence of Cps2E. (C) The extracted Cps2E glycolipid product (Und-P-P-[^{14}C]Glc) was incubated with membranes containing non-His-tagged Cps2T or membranes from the vector-only *E. coli* strain (MC). Reaction mixtures contained 3 μg of total membrane protein, 0.008% NP-40, 10 mM MnCl2, 1 mM DTT, 25 μl of the Und-P-P-[^{14}C]Glc solution (see Materials and Methods), and 0.1mM dTDP-Rha. Reactions were processed as described above. (D) Cps2E immunoblot demonstrating controls for extracted glycolipids. Lanes: 1, 3 μg of total membrane protein before extraction; 2, extracted organic phase containing glycolipids; 3, interface between organic and aqueous phase containing extracted proteins.

This product was not observed when membranes containing Cps2E were absent from the reactions (data not shown), suggesting that only Cps2E was capable of adding Glc to a lipid acceptor. When UDP-[^{14}C]Glc and dTDP-Rha were added to membranes that contained Cps2E and each of the other glycosyltransferases, only the reaction that contained Cps2T resulted in the formation of a new [^{14}C]Glc-labeled product (Fig. 6B). The slower migration of this product relative to [^{14}C]Glc is consistent with the addition of

a sugar residue(s) (34). Formation of this product only upon addition of dTDP-Rha and Cps2T indicates that Cps2T is the β1-4 rhamnosyltransferase that adds the second sugar to the repeat unit. Because both Cps2E and Cps2T were present in the glycosyltransferase reactions, we next tested whether Cps2T could add Rha to Und-P-P-Glc in the absence of Cps2E. Here, membranes containing Cps2E were used to synthesize Und-P-P-[^{14}C]Glc, which was then extracted and incubated with dTDP-Rha and membranes containing Cps2T. As shown in Fig. 6C, Cps2T was capable of adding Rha to Und-P-P-[^{14}C]Glc independently of Cps2E, which was undetectable in the glycolipid fraction by immunoblotting (Fig. 6D). The formation of this product in the absence of free UDP-[^{14}C]Glc further confirms the Cps2T addition of Rha to a Cps2E-dependent glycolipid.

DISCUSSION

Bacterial polysaccharide capsules are commonly assembled by the Wzy-dependent mechanism, where multiple glycosyltransferases synthesize the capsule repeat unit on a lipid acceptor. *cps2T* was first predicted to encode a rhamnosyltransferase by Iannelli *et al*. (32). Later bioinformatic analyses showed that each of the 14 solved capsule structures for serotypes containing a *cpsT* homologue has a β1-4 Rha at the second sugar position, leading to the prediction that Cps2T (WchF) is a β1-4 rhamnosyltransferase (1,39). Each of these serotypes also has a *cps2E* homologue, and as a result, the first two sugars and linkages in their capsule repeat unit backbone (Rhaβ1-4Glcβ1-) are conserved. The serotypes differ, however, in the remaining composition of sugars and linkages. All but two (serotypes 2 and 7B) lack a Rha residue at the third sugar position, strongly suggesting that Cps2T adds only a single Rha residue. In the type

2 capsule structure, two identical inverting linkages (α1-3Rha- α1-3 Rha) are present, but the *cps2* locus contains only one gene for an inverting transferase (*cps2F*, also annotated *wchG*). Biochemical data presented here confirm the assignment of Cps2T as the β1-4 rhamnosyltransferase that adds the first Rha, and genetic evidence discussed below indicates that Cps2F adds one or both of the subsequent Rha residues.

Knowledge of the mechanisms that regulate the repeat unit assembly process is limited, but it is a common property of many pathways that the committed step serves as an important point of regulation. The committed step is generally considered to reflect the first irreversible step in a pathway. In capsule synthesis, this step leads to the formation of a lipid-linked saccharide product, the accumulation of which can be lethal if the pathway is stalled prior to transfer of the sugars from Und-P to another acceptor. The lethality is thought to be due to the sequestering of the limited quantities of Und-P in the cell (41) that are required for other essential pathways or to the destabilization of the membrane (8, 14, 33, 49, 61, 64). Based on the results of a previous study, we anticipated that the committed step in type 2 capsule synthesis might not be the addition of the first Glc. There, mutations resulting in partial repeat unit formation due to a lack of the terminal GlcUA or mutations that inhibit repeat unit transport (Wzx flippase mutants) or polymerization (Wzy polymerase mutants) were obtained only in the presence of secondary suppressor mutations (61). Although most of these suppressor mutations mapped to *cps2E*, one was an insertion mutation in *cps2L* that was likely polar on *cps2M*, *cps2N*, and *cps2O*, thereby eliminating synthesis of dTDP-Rha and capsule. Cps2E activity from the latter strain was comparable to that of the parent in *in vitro* assays, and no mutations were present in *cps2E* or other capsule genes. The results of the present

study are consistent with the notion that the Cps2E catalyzed addition of the first sugar (Glc) is not the committed step, as elimination of Cps2T activity did not require suppressor mutations. Cps2E activity *in vivo* therefore either is reversible (as has been demonstrated for the *in vitro* reaction [12]) or requires the presence of the subsequent substrate and/or enzyme to proceed efficiently. An alternate hypothesis is that Und-P-P-Glc is not toxic because it can be used in another pathway, such as glycoprotein synthesis (31). However, if Und-P-P-Glc does occur in other *S. pneumoniae* pathways, it is unavailable for use in capsule synthesis, as Δ*cps2E* mutants do not produce CPS. The presence of secondary mutations in only the Δ*cps2F*, Δ*cps2G*, and Δ*cps2I* mutants provides genetic evidence that Cps2F, Cps2G, and Cps2I act downstream of the committed step and Cps2T performs this step.

A requirement for suppressor mutations to prevent lethality resulting from mutations affecting later-acting steps has similarly been reported for lipopolysaccharide O-antigen and exopolysaccharide mutants in Gram-negative bacteria (8, 33, 49, 64) and teichoic acid mutants in the Gram-positive bacterium *Staphylococcus aureus* (14). The *S. pneumoniae* capsule system seems to differ from some of these systems, however, in that addition of the second sugar can be eliminated without apparent detriment. It remains unclear why most secondary mutations occur in *cps2E* (30 among 33 mutants in our studies to date), unless it provides a target larger and more effective than the genes for other enzymes. In both this and our previous study (61), mutations in either the Cps2E cytoplasmic loop, which encodes the glycosyltransferase activity, or extracytoplasmic loop were effective in suppressing the effects of the primary mutation. Both domains thus

provide important functions, though the role of the extracytoplasmic loop remains unknown.

Deletion of *cps2G* or *cps2I* required suppressor mutations for viable mutants, and the observation that these mutants were acquired at a low frequency supports the notion that these gene products are essential in the parental background. Despite the requirement for a suppressor mutation, Δ*cps2F* mutants were obtained at a higher frequency than Δ*cps2G* or Δ*cps2I* mutants and at a frequency more similar to that for Δ*cps2T* mutants. This result suggests that these suppressor mutations arose at a different point in time than those in the rarely obtained Δ*cps2G* and Δ*cps2I* mutants and may reflect a different stringency for tolerance of the initial mutation. If loss of *cps2G* or *cps2I* is rapidly lethal, viability of these mutants may require the initial transformation to occur in those rare recipients that already contain a *cps2E* mutation. In contrast, if Δ*cps2F* mutants are viable for many generations, the deletion mutation could be obtained at a high frequency and the necessary *cps2E* mutation could occur in any of the transformants over many generations. Scenarios consistent with this possibility are a less efficient Cps2T when Cps2F is missing and/or feedback inhibition of the dTDP-Rha biosynthetic enzymes (40) as dTDP-Rha accumulates in the Δ*cps2F* mutants. Either effect would slow the accumulation of Und-P-P-Glc-Rha compared to that in the Δ*cps2G* and Δ*cps2I* mutants, with a consequent reduction in time to death. These results are also consistent with Cps2F activity immediately succeeding that of Cps2T, as suggested by the assignment of Cps2F as the sole inverting transferase in the type 2 locus (1, 39). Because the type 2 capsule has two identical inverting linkages (α 1-3 Rha-α 1-3 Rha), Cps2F may catalyze two sequential reactions, a dual functionality that was first described in eukaryotes (18) but

only recently characterized for prokaryotic glycosyltransferases (51). We are currently testing this possibility, in addition to identifying the functions of the other glycosyltransferases.

Polymerization of the Und-P-P-Glc-Rha-Rha-Rha-Glc-GlcUA repeat unit results in the formation of a Glc-GlcUA side chain in the type 2 polymer (Fig. 1). Detection of a small amount of CPS in the Δ*cps2G* mutants must mean that the *cps2E* suppressor mutations in these strains do not completely abolish Cps2E activity *in vivo*, resulting in sublethal amounts of partial repeat units that can be translocated and polymerized.Asimilar capsule phenotype was observed with the Δ*cps2K* UDP-Glc dehydrogenase mutants, which also contained suppressor mutations that primarily mapped to *cps2E* (61). These results suggest that *cps2G* encodes the glycosyltransferase that adds the Glc or GlcUA, and its elimination does not preclude synthesis of the backbone. Cps2I is expected to add the other sugar in the resulting side chain, but we have not been able to detect any polymer with the Δ*cps2I* mutants. It is not yet known whether this result is due to its true absence, to polymer production below our limit of detection, or to an inactive Cps2E in each of the mutants obtained to date. Capsule transfer to the cell wall in the Δ*cps2G* and Δ*cps2K* mutants has also not been detected (61; this study). Although this result may be due to detection limits, failure to complete capsule transfer to the cell wall could contribute to the lethality that requires a suppressor mutation for survival. Reduced efficiencies of the Wzx flippase and Wzy polymerase in recognizing the incomplete repeat units may also be important, as the Wzx flippase for *Salmonella enterica* group B and D2 Oantigen demonstrates a higher efficiency of translocation for the complete repeat unit than for one lacking a side-branch sugar (30).

When expressed in *E. coli*, Cps2T, as well as Cps2F, Cps2G, and Cps2I, behaved as a peripheral membrane protein. Glycosyltransferases responsible for synthesizing O antigens and other polysaccharide capsules have similarly been reported to localize to the membrane (48). This location is likely important in their ability to interact with the growing lipid-linked repeat unit and possibly with the other glycosyltransferases. *In vitro*, Cps2T catalyzed the formation of Und-P-P-Glc-Rha from Und-P-P-Glc and dTDP-Rha in the absence of Cps2E. The rhamnosyltransferase activity of Cps2T therefore does not require interaction with Cps2E, but we cannot rule out the possibility that interaction enhances *in vivo* activity and repeat unit formation, as suggested above.

Homologues of Cps2T are present in at least 27 *S. pneumoniae* serotypes (1, 5, 39). This abundance elevates *cpsT* to one of the most prevalent *S. pneumoniae* glycosyltransferase-encoding genes in the *cps* locus, second only to *cpsE* homologues. As a result of this work, we predict that Cps2T homologues in each of these *S. pneumoniae* serotypes are responsible for the committed step in capsule synthesis. The addition of Rha to Und-P-P-Glc as the committed step raises the possibility that regulation of Cps2T rhamnosyltransferase activity is a critical point of control. A principal factor regulating the action of glycosyltransferases is the availability of their nucleotide-sugar donor (54). The synthesis of dTDP-Rha is resource intensive, requiring the activity of four enzymes and two cofactors (Fig. 4A). Biochemical studies by Melo and Glaser (40) demonstrated that the dTDP-Rha biosynthetic pathway in *P. aeruginosa* possesses a feedback regulation where the end product, dTDP-Rha, competitively and noncompetitively inhibits the activity of the first enzyme in the pathway, which is a Cps2L homologue. These enzymes are also pyrophosphorylases, capable of converting dTDP-Glc to dTTP

and Glc-1-P (Fig. 4A) (40). Cps2L could therefore be an important cellular target for regulating dTDP-Rha concentrations, Cps2T activity, and ultimately, capsule levels. Interestingly, recent findings in *Caulobacter crescentus* have identified homologues of Cps2L and Cps2M (the third enzyme in the dTDP-Rha biosynthesis pathway) to be phosphorylated at serine and/or threonine residues in response to carbon starvation and rich environments, respectively (26). If Cps2L and/or Cps2M is phosphorylated in *S. pneumoniae*, these modifications could affect dTDP-Rha synthesis and, correspondingly, capsule levels. Homologues of *cpsLNMO* are present in at least 38 *S. pneumoniae* serotypes (5). *S. pneumoniae* possesses a tyrosine phosphorylation system, encoded by *cpsBCD*, which is known to influence capsule levels but through an unidentified mechanism (4, 43, 58). Regulation of nucleotide-sugars as a modulator of capsule levels has been observed in *S. pneumoniae* serotype 3 strains that utilize a synthase-dependent mechanism for capsule synthesis (17, 55) but thus far has not been described in Wzy dependent synthesis.

ACKNOWLEDGMENTS

This work was supported by grants to J.Y. from the National Institutes of Health (AI28457), the Mizutani Foundation for Glycoscience, and the UAB Health Services Foundation. Funds for the operation of the Targeted Metabolomics and Proteomics (TMP) core facility come in part from the UAB Center for Nutrient-Gene Interaction (U54 CA 100949), the Purdue- UAB Botanicals Center for Age-Related Disease (P50 AT00477), the UAB O'Brien Acute Kidney Injury Center (P30 DK079337), the UAB Skin Disease Research Center (P30 AR50948), and the UAB Lung Health Center. DNA

sequencing was performed by the UAB CFAR DNA sequencing core (P30A127767) or the Genomics Core Facility of the Heflin Center for Genomic Science.

We thank Doyle Ray Moore II from the University of Alabama at Birmingham TMP core facility for assistance with ESI-MS/MS analyses and Kellie Putnam for generating *E. coli* strain KJ4152.

REFERENCES

1. **Aanensen, D. M., A. Mavroidi, S. D. Bentley, P. R. Reeves, and B. G. Spratt.** 2007. Predicted functions and linkage specificities of the products of the *Streptococcus pneumoniae* capsular biosynthetic loci. J Bacteriol **189:**7856-76.
2. **Avery, O. T., and R. Dubos.** 1931. The protective action of a specific enzyme against type III pneumococcus infection in mice. J Exp Med **54:**73-89.
3. **Avery, O. T., C. M. MacLeod, and M. McCarty.** 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:**137-58.
4. **Bender, M. H., R. T. Cartee, and J. Yother.** 2003. Positive correlation between tyrosine phosphorylation of CpsD and capsular polysaccharide production in *Streptococcus pneumoniae*. J Bacteriol **185:**6057-66.
5. **Bentley, S. D., D. M. Aanensen, A. Mavroidi, D. Saunders, E. Rabbinowitsch, M. Collins, K. Donohoe, D. Harris, L. Murphy, M. A. Quail, G. Samuel, I. C. Skovsted, M. S. Kaltoft, B. Barrell, P. R. Reeves, J. Parkhill, and B. G. Spratt.** 2006. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. PLoS Genet **2:**e31.
6. **Breton, C., L. Snajdrova, C. Jeanneau, J. Koca, and A. Imberty.** 2006. Structures and mechanisms of glycosyltransferases. Glycobiology **16:**29R-37R.
7. **Brown, E. J., K. A. Joiner, R. M. Cole, and M. Berger.** 1983. Localization of complement component 3 on *Streptococcus pneumoniae*: anti-capsular antibody causes complement deposition on the pneumococcal capsule. Infect Immun **39:**403-9.
8. **Burrows, L. L., and J. S. Lam.** 1999. Effect of *wzx* (*rfbX*) mutations on A-band and B-band lipopolysaccharide biosynthesis in *Pseudomonas aeruginosa* O5. J Bacteriol **181:**973-80.
9. **Calix, J. J., and M. H. Nahm.** 2010. A new pneumococcal serotype, 11E, has a variably inactivated *wcjE* gene. J Infect Dis **202:**29-38.
10. **Caputto, R., L. F. Leloir, C. E. Cardini, and A. C. Paladini.** 1950. Isolation of the coenzyme of the galactose phosphate-glucose phosphate transformation. J Biol Chem **184:**333-50.

11. **Cardini, C. E., A. C. Paladini, R. Caputto, and L. F. Leloir.** 1950. Uridine diphosphate glucose: The coenzyme of the galactose-glucose phosphate isomerization. Nature **165:**191.
12. **Cartee, R. T., W. T. Forsee, M. H. Bender, K. D. Ambrose, and J. Yother.** 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:**7425-33.
13. **Centers for Disease Control and Prevention.** Pneumococcal Disease, Epidemiology and Prevention of Vaccine-Preventable Diseases. Atkinson W, Wolfe S, Hamborsky J, eds., 12th ed. Washington DC: Public Health Foundation.2011
14. **D'Elia, M. A., M. P. Pereira, Y. S. Chung, W. Zhao, A. Chau, T. J. Kenney, M. C. Sulavik, T. A. Black, and E. D. Brown.** 2006. Lesions in teichoic acid biosynthesis in *Staphylococcus aureus* lead to a lethal gain of function in the otherwise dispensable pathway. J Bacteriol **188:**4183-9.
15. **Dillard, J. P., M. W. Vandersea, and J. Yother.** 1995. Characterization of the cassette containing genes for type 3 capsular polysaccharide biosynthesis in *Streptococcus pneumoniae*. J Exp Med **181:**973-83.
16. **Dong, S., S. A. McPherson, Y. Wang, M. Li, P. Wang, C. L. Turnbough, Jr., and D. G. Pritchard.** 2010. Characterization of the enzymes encoded by the anthrose biosynthetic operon of *Bacillus anthracis*. J Bacteriol **192:**5053-62.
17. **Forsee, W. T., R. T. Cartee, and J. Yother.** 2009. A kinetic model for chain length modulation of *Streptococcus pneumoniae* cellubiuronan capsular polysaccharide by nucleotide sugar donor concentrations. J Biol Chem **284:**11836-44.
18. **Frank, C. G., and M. Aebi.** 2005. ALG9 mannosyltransferase is involved in two different steps of lipid-linked oligosaccharide biosynthesis. Glycobiology **15:**1156-63.
19. **Fujiki, Y., A. L. Hubbard, S. Fowler, and P. B. Lazarow.** 1982. Isolation of intracellular membranes by means of sodium carbonate treatment: application to endoplasmic reticulum. J Cell Biol **93:**97-102.
20. **Giraud, M. F., and J. H. Naismith.** 2000. The rhamnose pathway. Curr Opin Struct Biol **10:**687-96.
21. **Glaser, L., and S. Kornfeld.** 1961. The enzymatic synthesis of thymidine-linked sugars. II. Thymidine diphosphate L-rhamnose. J Biol Chem **236:**1795-9.
22. **Graninger, M., B. Kneidinger, K. Bruno, A. Scheberl, and P. Messner.** 2002. Homologs of the Rml enzymes from *Salmonella enterica* are responsible for dTDP-beta-L-rhamnose biosynthesis in the gram-positive thermophile *Aneurinibacillus thermoaerophilus* DSM 10155. Appl Environ Microbiol **68:**3708-15.
23. **Graninger, M., B. Nidetzky, D. E. Heinrichs, C. Whitfield, and P. Messner.** 1999. Characterization of dTDP-4-dehydrorhamnose 3,5-epimerase and dTDP-4-dehydrorhamnose reductase, required for dTDP-L-rhamnose biosynthesis in *Salmonella enterica* serovar Typhimurium LT2. J Biol Chem **274:**25069-77.
24. **Griffith, F.** 1928. The Significance of Pneumococcal Types. J Hyg (Lond) **27:**113-59.

25. **Guo, H., W. Yi, J. K. Song, and P. G. Wang.** 2008. Current understanding on biosynthesis of microbial polysaccharides. Curr Top Med Chem **8:**141-51.
26. **Ham, B. M.** 2011. Prokaryotic phosphorylation of serine, threonine, and tyrosine, p 181-244. Proteomics of Biological Systems. John Wiley & Sons, Inc.
27. **Hardy, G. G., M. J. Caimano, and J. Yother.** 2000. Capsule biosynthesis and basic metabolism in *Streptococcus pneumoniae* are linked through the cellular phosphoglucomutase. J Bacteriol **182:**1854-63.
28. **Hardy, G. G., A. D. Magee, C. L. Ventura, M. J. Caimano, and J. Yother.** 2001. Essential role for cellular phosphoglucomutase in virulence of type 3 *Streptococcus pneumoniae*. Infect Immun **69:**2309-17.
29. **Henrichsen, J.** 1995. Six newly recognized types of *Streptococcus pneumoniae*. J Clin Microbiol **33:**2759-62.
30. **Hong, Y., M. M. Cunneen, and P. R. Reeves.** 2012. The Wzx translocases for *Salmonella enterica* O-antigen processing have unexpected serotype specificity. Mol Microbiol. doi: 10.1111/j.1365-2958.2012.08048.x. Epub 2012 Apr 13.
31. **Hug, I., and M. F. Feldman.** 2011. Analogies and homologies in lipopolysaccharide and glycoprotein biosynthesis in bacteria. Glycobiology **21:**138-51.
32. **Iannelli, F., B. J. Pearce, and G. Pozzi.** 1999. The type 2 capsule locus of *Streptococcus pneumoniae*. J Bacteriol **181:**2652-4.
33. **Katzen, F., D. U. Ferreiro, C. G. Oddo, M. V. Ielmini, A. Becker, A. Puhler, and L. Ielpi.** 1998. *Xanthomonas campestris* pv. campestris gum mutants: effects on xanthan biosynthesis and plant virulence. J Bacteriol **180:**1607-17.
34. **Kolkman, M. A., B. A. van der Zeijst, and P. J. Nuijten.** 1997. Functional analysis of glycosyltransferases encoded by the capsular polysaccharide biosynthesis locus of *Streptococcus pneumoniae* serotype 14. J Biol Chem **272:**19502-8.
35. **Kornfeld, S., and L. Glaser.** 1961. The enzymic synthesis of thymidine-linked sugars. I. Thymidine diphosphate glucose. J Biol Chem **236:**1791-4.
36. **Lairson, L. L., B. Henrissat, G. J. Davies, and S. G. Withers.** 2008. Glycosyltransferases: structures, functions, and mechanisms. Annu Rev Biochem **77:**521-55.
37. **Llull, D., E. Garcia, and R. Lopez.** 2001. Tts, a processive beta-glucosyltransferase of *Streptococcus pneumoniae*, directs the synthesis of the branched type 37 capsular polysaccharide in pneumococcus and other gram-positive species. J Biol Chem **276:**21053-61.
38. **Magee, A. D., and J. Yother.** 2001. Requirement for capsule in colonization by *Streptococcus pneumoniae*. Infect Immun **69:**3755-61.
39. **Mavroidi, A., D. M. Aanensen, D. Godoy, I. C. Skovsted, M. S. Kaltoft, P. R. Reeves, S. D. Bentley, and B. G. Spratt.** 2007. Genetic relatedness of the *Streptococcus pneumoniae* capsular biosynthetic loci. J Bacteriol **189:**7841-55.
40. **Melo, A., and L. Glaser.** 1965. The nucleotide specificity and feedback control of thymidine diphosphate D-glucose pyrophosphorylase. J Biol Chem **240:**398-405.
41. **Mengin-Lecreulx, D., L. Texier, M. Rousseau, and J. van Heijenoort.** 1991. The *murG* gene of *Escherichia coli* codes for the UDP-N-acetylglucosamine: N-

acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase involved in the membrane steps of peptidoglycan synthesis. J Bacteriol **173:**4625-36.

42. **Morona, J. K., R. Morona, and J. C. Paton.** 1997. Characterization of the locus encoding the *Streptococcus pneumoniae* type 19F capsular polysaccharide biosynthetic pathway. Mol Microbiol **23:**751-63.
43. **Morona, J. K., J. C. Paton, D. C. Miller, and R. Morona.** 2000. Tyrosine phosphorylation of CpsD negatively regulates capsular polysaccharide biosynthesis in *Streptococcus pneumoniae*. Mol Microbiol **35:**1431-42.
44. **Nelson, A. L., A. M. Roche, J. M. Gould, K. Chim, A. J. Ratner, and J. N. Weiser.** 2007. Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. Infect Immun **75:**83-90.
45. **Neufeld, F.** 1902. Ueber die Agglutination der Pneumokokken und über die Theorieen der Agglutination. Medical Microbiology and Immunology **40:**54.
46. **Okazaki, R., Okazakit, J. L. Strominger, and A. M. Michelson.** 1962. Thymidine diphosphate 4-keto-6-deoxy-d-glucose, an intermediate in thymidine diphosphate L-rhamnose synthesis in *Escherichia coli* strains. J Biol Chem **237:**3014-26.
47. **Pelosi, L., M. Boumedienne, N. Saksouk, J. Geiselmann, and R. A. Geremia.** 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:**857-65.
48. **Raetz, C. R., and C. Whitfield.** 2002. Lipopolysaccharide endotoxins. Annu Rev Biochem **71:**635-700.
49. **Rick, P. D., K. Barr, K. Sankaran, J. Kajimura, J. S. Rush, and C. J. Waechter.** 2003. Evidence that the *wzxE* gene of *Escherichia coli* K-12 encodes a protein involved in the transbilayer movement of a trisaccharide-lipid intermediate in the assembly of enterobacterial common antigen. J Biol Chem **278:**16534-42.
50. **Stern, R. J., T. Y. Lee, T. J. Lee, W. Yan, M. S. Scherman, V. D. Vissa, S. K. Kim, B. L. Wanner, and M. R. McNeil.** 1999. Conversion of dTDP-4-keto-6-deoxyglucose to free dTDP-4-keto-rhamnose by the rmIC gene products of Escherichia coli and Mycobacterium tuberculosis. Microbiology **145 (Pt 3):**663-71.
51. **Troutman, J. M., and B. Imperiali.** 2009. *Campylobacter jejuni* PglH is a single active site processive polymerase that utilizes product inhibition to limit sequential glycosyl transfer reactions. Biochemistry **48:**2807-16.
52. **Valvano, M. A.** 2003. Export of O-specific lipopolysaccharide. Front Biosci **8:**s452-71.
53. **van Dam, J. E., A. Fleer, and H. Snippe.** 1990. Immunogenicity and immunochemistry of Streptococcus pneumoniae capsular polysaccharides. Antonie Van Leeuwenhoek **58:**1-47.
54. **Varki A., C. R., Esko J., Freeze H., Hart G., and J. Marth.** 1999. Glycosyltransferases, 1st ed.: Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY).

55. **Ventura, C. L., R. T. Cartee, W. T. Forsee, and J. Yother.** 2006. Control of capsular polysaccharide chain length by UDP-sugar substrate concentrations in *Streptococcus pneumoniae*. Mol Microbiol **61:**723-33.
56. **Weadge, J. T., M. M. Palcic, and T. P. Begley.** 2007. Glycosyltransferases, Chemistry of. John Wiley & Sons, Inc.
57. **Weiser, J. N., R. Austrian, P. K. Sreenivasan, and H. R. Masure.** 1994. Phase variation in pneumococcal opacity: relationship between colonial morphology and nasopharyngeal colonization. Infect Immun **62:**2582-9.
58. **Weiser, J. N., D. Bae, H. Epino, S. B. Gordon, M. Kapoor, L. A. Zenewicz, and M. Shchepetov.** 2001. Changes in availability of oxygen accentuate differences in capsular polysaccharide expression by phenotypic variants and clinical isolates of *Streptococcus pneumoniae*. Infect Immun **69:**5430-9.
59. **Whitfield, C.** 2006. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. Annu Rev Biochem **75:**39-68.
60. **Wood, W. B., Jr., and M. R. Smith.** 1949. The inhibition of surface phagocytosis by the capsular slime layer of pneumococcus type III. J Exp Med **90:**85-96.
61. **Xayarath, B., and J. Yother.** 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:**3369-81.
62. **Yother, J.** 2011. Capsules of *Streptococcus pneumoniae* and other bacteria: paradigms for polysaccharide biosynthesis and regulation. Annu Rev Microbiol **65:**563-81.
63. **Yother, J.** 2006. Integration of capsular polysaccharide biosynthesis with metabolic and virulence pathways in *Streptococcus pneumoniae*, pp 51-65. *In* Brogden, K. A. et. al (ed), Virulence mechanisms of bacterial pathogens 4th ed. ASM Press, Washington, D.C.
64. **Yuasa, R., M. Levinthal, and H. Nikaido.** 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:**433-44.

STREPTOCOCCUS PNEUMONIAE SEROTYPE 2 CAPSULAR GLYCOSYLTRANSFERASES: BIOCHEMICAL ACTIVITIES AND SIGNIFICANCE OF SUPPRESSOR MUTATIONS

DAVID B.A. JAMES AND JANET YOTHER

In preparation for *Journal of Bacteriology*

Format adapted for dissertation

ABSTRACT

The capsule 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 *S. pneumoniae* serotype 2, the majority of these secondary mutations map to the gene encoding the initiating capsular glycosyltransferase (*cps2E*). Here, we provide a consensus amino acid sequence that reflects the 69 *S. pneumoniae* serotypes containing Cps2E homologues and demonstrate that there is a highly significant association between the residues that are 100% conserved and those selected as suppressor mutations. Cps2E contains an extracytoplasmic loop whose function is unknown. By analyzing our collection of mutants, we identified 5 amino acids (A162, D167, V196, L200, H258) selected as suppressor mutations in the extracytoplasmic domain of Cps2E, and demonstrate that these residues are highly conserved among *S. pneumoniae* serotypes and required for efficient cytoplasmic sugar transfer *in vitro*. The critical function of these amino acids and their extracytoplasmic location may suggest an important point of contact with other enzymes/environmental factors. In addition, we also clarify here the function of three glycosyltransferases (Cps2F, Cps2G and Cps2I) whose deletions also selected for secondary mutations in *cps2E*. We utilized *in vitro* glycosyltransferase assays to demonstrate that Cps2F catalyzes the third and fourth linkage, while Cps2G catalyzes the fifth linkage in the capsule repeat unit. Addition of the terminal sugar residue could not be detected, however activities of the other glycosyltransferases together with bioinformatic analyses suggest this step is mediated by Cps2I.

INTRODUCTION

Streptococcus pneumoniae (the pneumococcus) is an extracellular bacterial pathogen that produces a protective surface structure known as the capsular polysaccharide (CPS) (28). The CPS is an essential component for virulence, influencing both nasopharyngeal colonization and systemic infection (2, 5, 14, 15, 24, 25, 39, 41). CPS is also an immunodominant surface antigen, serving as the basis for *S. pneumoniae* serological typing (13, 14, 23, 35). Over 90 serotypes have been identified and all but two utilize the Wzy-dependent pathway to synthesize CPS, while the remaining serotypes (serotypes 3 and 37) utilize the synthase-dependent pathway (4, 10, 22, 43). This study focuses on the Wzy-dependent pathway which was first identified as the mechanism of synthesis for lipopolysaccharide O-antigen in Gram-negative bacteria (29, 34, 40), but is the principle mechanism for CPS synthesis in Gram-positive bacteria (43).

In the Wzy-dependent pathway the capsule repeat unit is synthesized on a membrane acceptor, undecaprenyl phosphate (Und-P), by multiple glycosyltransferases. Once complete, 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 serotype 2 which has a CPS composed of three sugar residues: glucose (Glc), rhamnose (Rha) and glucuronic acid (GlcUA) (18).

Our laboratory has previously identified and characterized the first two enzymes involved in synthesizing the serotype 2 repeat unit; the initiating glucose-1-phosphate (Glc-1-P) transferase Cps2E (8), and the β1-4 rhamnosyltransferase Cps2T which catalyzes the second and committed step (17). Three other genes in the *cps2* locus, *cps2F*,

cps2G and *cps2I*, are predicted to encode the glycosyltransferases required for complete synthesis of the repeat unit but biochemical evidence supporting their functions is lacking (1, 16, 17) (Fig. 1).

There are a limited number of publications characterizing pneumococcal glycosyltransferases involved in Wzy-dependent capsule synthesis. The initiating glycosyltransferase CpsE and its homologues are well characterized, with analysis of enzyme function demonstrated in serotypes 2, 8, 9v and 14 (8, 20, 27, 36). CpsE is a member of the polyprenyl-phosphate hexose-1-phosphate transferase (PHPT) family (34), where *Salmonella enterica* galactose-1-phosphate transferase WbaP serves as the prototypic enzyme (19, 44).

Three functional domains are present in PHPT enzymes; the N terminal domain, comprising 4 membrane spanning regions, a central domain containing an extracytoplasmic loop, and the C terminal domain comprising 1 membrane spanning region and a cytoplasmic tail (30). The C terminal domain catalyzes sugar transfer, as independent studies have described this domain to be sufficient for glycosyltransferase activity (26, 27, 30, 37). The N terminal domain is suggested to be required for protein insertion or stability in the membrane (30). The function of the extracytoplasmic loop is less clear. This domain is speculated to be involved in release of UndP-P-linked sugar from the protein (37), while others have suggested a role in modulation of the chain-length distribution (30). At least 69 *S. pneumoniae* serotypes express homologues of Cps2E (also known as WchA) (4).

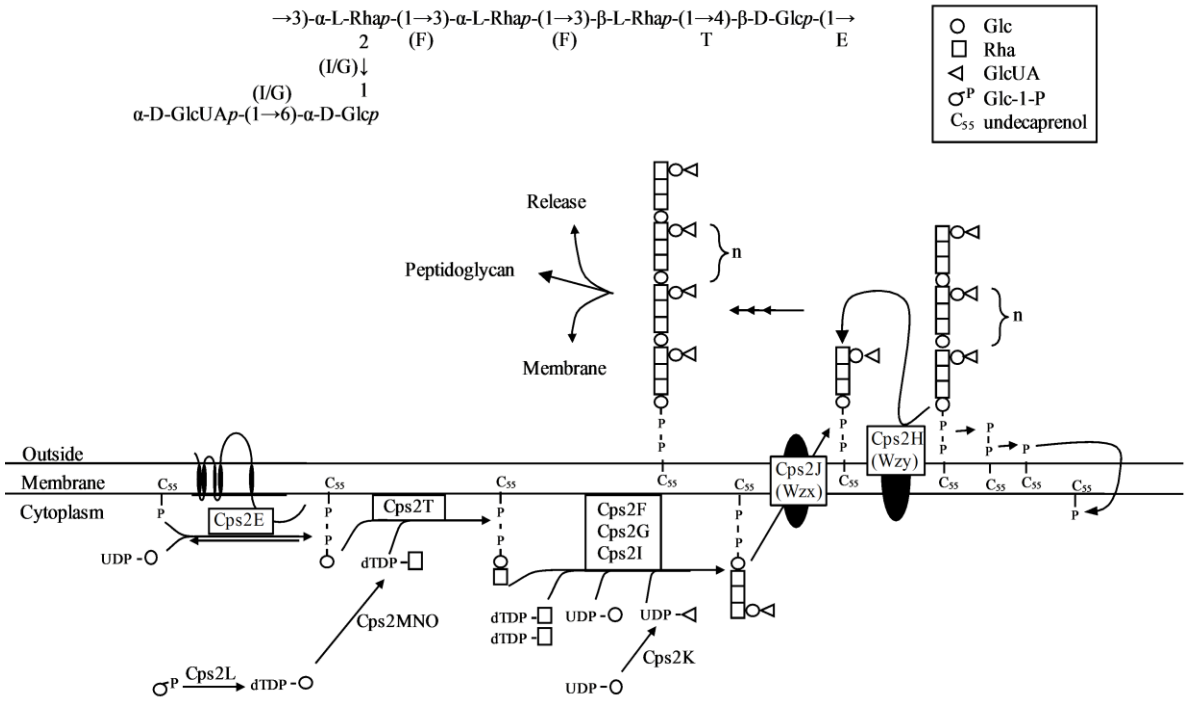

FIG 1. *S. pneumoniae* type 2 capsular polysaccharide repeat unit structure and Wzy-dependent pathway. (Top) Repeat unit structure. Letters below each glycosidic linkage represent the glycosyltransferase known or predicted to catalyze the linkage. Predicted enzymes are in parenthesis and are based on reference 1, where the specificities of Cps2G and Cps2I were unclear. Glc, glucose; Rha, rhamnose; GlcUA, glucuronic acid. E, Cps2E; T, Cps2T; F, Cps2F; G, Cps2G; I, Cps2I. (Bottom) Wzy-dependent model for synthesis of the type 2 repeat unit. Synthesis of the capsule repeat unit occurs through the activity of multiple glycosyltransferases. The completed repeat unit is transferred to the outer face of the cytoplasmic membrane, polymerized and then remains associated with the membrane, released or attached to peptidoglycan. Adapted from reference (63). Used with permission of ASM Press.

In previous reports, we demonstrated that deletion of *cps2* genes encoding glycosyltransferases (*cps2F*, *cps2G*, *cps2I*), a UDP-glucose dehydrogenase (*cps2K*), the Wzx translocase (*cps2J*) or the Wzy polymerase (*cps2H*) were lethal (17, 42). Mutants were selected only in the presence of secondary mutations, the majority of which map to *cps2E* (17, 42). The lethality was attributed to either destabilization of the membrane by incomplete repeat units or sequestration of Und-P in the capsule pathway and away from

essential pathways such as peptidoglycan synthesis. To date, we have obtained over 30 suppressor mutations in *cps2E*. All of the mutations are unique and observed in the C terminal domain as well as in the central extracytoplasmic domain (17, 42).

Suppressor mutations mapping to the C terminal domain can be interpreted as a direct inhibition of Glc-1-P transferase activity, as this domain alone contains glycosyltransferase 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 utilize our novel collection of secondary mutations in *cps2E* to investigate the function of the extracytoplasmic loop domain. We demonstrate that these residues are critical for capsule initiation and thus provide additional rational for crystallization studies. We also provide biochemical evidence for other previously uncharacterized capsular glycosyltransferases which, together with bioinformatic analyses, completes the transferase specificity for the complete type 2 CPS repeat unit in *S. pneumoniae*.

MATERIALS AND METHODS

Materials. UDP- [^{14}C]Glc (300 mCi/mmol) was obtained from American Radiochemical Chemicals and UDP- [^{3}H]Glc (43.8 Ci/mmol) was obtained from Perkin Elmer. Inorganic pyrophosphatase, NAD^{+}, NADH, dTTP, and Glc-1-P were all obtained from Sigma-Aldrich. Nonident P-40 (NP-40) was obtained from Pierce. Solid-phase extraction (SPE) carbograph (150 mg/4 ml) columns were obtained from Grace Davison Discovery Sciences.

Bacterial growth conditions. Broth and plate cultures of *S. pneumoniae* and *E. coli* (Table 1)were grown as previously described (17). Media were supplemented with the following antibiotics when 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).

Glycosyltransferase expression and isolation. Cultures expressing Cps2E, Cps2T, Cps2F, Cps2G and Cps2I were processed as previously described (17). Briefly, 250 ml cultures were grown under the appropriate antibiotic selection to mid-exponential phase (cell density of approximately 4.5 x 10^8 CFU/ml) and then induced with 0.2% arabinose .Cells were centrifuged, washed, lysed and membranes were collected as previously described (17), with the exception that membranes were centrifuged at 100,000 x *g* instead of 50,000 x *g*. The final membrane pellet was suspended in 500 μl of 100 mM Tris-Acetate (pH 7.5) containing 10% glycerol and 1 mM DTT. Total membrane proteins were quantified using the Bradford Bio-Rad protein dye assay.

Synthesis, purification and quantification of dTDP-Rha. dTDP-Rha was synthesized in two steps essentially as previously described (17). Briefly, dTDP-Glc was synthesized in a 350 μl reaction mixture containing 50 mM HEPES pH 7.6, 45 mM $MgCl_2$, 3.4 mM dTTP, 3.4 mM Glc-1-P, 2 units of inorganic pyrophosphatase and 2 μM Cps2L. After an incubation of 90 min at 37^{o}C, the proteins were removed by Amicon Ultra 10 kDa centrifugal filters (Millipore).

TABLE 1. *S. pneumoniae* and *E. coli* strains used in this study

Strain(s)	Properties[a]	References or source
S. pneumoniae strains		
BX547	D39 Δ*cps2K*, Cps^{r}; *cps2E* $^{5544\ G\rightarrow T}$ (D167Y) ; Km^{R}	(42)
BX552	D39 Δ*cps2H*, Cps^{-}; *cps2E* $^{5539\ C\rightarrow T}$ (L200F) ; Km^{R}	(42)
BX605	D39 Δ*cps2K*, Cps^{r}; *cps2E* $^{5632\ T\rightarrow G}$ (V196G) ; Km^{R}	(42)
D39	Type 2 parent strain, Cps^{+}	(3)
DJ916	D39 *Δcps2F*, Cps^{-} ; *cps2E* $^{5718\ A\rightarrow G}$ (H258R); Km^{R}	(17)
DJ921	D39 *Δcps2G*, Cps^{r} ; *cps2E* $^{5529G\rightarrow T}$ (A162S); Km^{R}	(17)
KA1522	D39 *Δcps2E*, Cps^{-}	(8)
E. coli strains		
BL21-AI	F^{-} *ompT hsdSB*(r_B^{-} m_B^{-}) *gal dcm araB::*T7 RNAP-*tetA*	Invitrogen
DJ005	BL21-AI expressing full length Cps2I; Ap^{R}	(17)
DJ009	BL21-AI expressing full length Cps2T; Ap^{R}	(17)
DJ011	BL21-AI expressing full length Cps2G; Ap^{R}	(17)
DJ204	BL21-AI expressing full length Cps2F, replaced GTG start with ATG; Ap^{R}	(17)
KJ4152	BL21-AI expressing full length Cps2E; Ap^{R}	(17)
RC124	BL21-AI vector control; Ap^{R}	(8)

[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; Km^{R}, kanamycin resistant; Ap^{R}, ampicillin resistant.

The filtrate was used in the second step and supplemented with 0.1 mM NAD^+, 6 mM NADH, 1 µM Cps2N, 1 µM Cps2M, and 1 µM Cps2O in a final volume of 350 µl. The reaction was incubated for 30 min at 37^{o}C, the proteins were filtered, and dTDP-Rha was separated by HPLC as described (17). Fractions containing dTDP-Rha were pooled, desalted, freeze-dried as described (17). For use in glycosyltransferase assays, samples were re-suspended in 100 mM Tris-Acetate pH 7.6 and dTDP-Rha concentrations were calculated from absorbances at 260 nm using Beer's law and the thymidine extinction coefficient (7.4 x 10^3 M^{-1} cm^{-1}).

Glycosyltansferase assays. Glycosyltansferase assays were performed as previously described (17). Briefly, membranes containing 3 µg of total protein were incubated for 1 h at 10^{o}C in reactions containing 10 mM $MnCl_2$, 1 mM DTT, 0.008% NP-40 (when utilizing multiple glycosyltransferases) , 0.025 µCi 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 µl with 100 mM Tris-Acetate pH 7.6. Reactions were stopped with 1 mL of 2:1 chroloform:methanol and the organic phase was extracted three times with pure solvent upper phase (PSUP; 1.5 mL chloroform, 25 mL methanol, 23.5 mL H_2O, and 0.183 g KCl). Glycolipids were either preserved for later use as described for Und-P-P-^{14}C-Glc-Rha-Rha-Rha (see below) or subjected to mild-acid hydrolysis by resuspending the dried organic phase in 80 µL of 20 mM HCL and incubating at 70^{o}C for 20 min.

Hydrolyzed samples were dried, resuspended in 10 µl 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 phosphor imager (General Electric Healthcare).

Synthesis of Und-P-P-[^{14}C] Glc-Rha-Rha-Rha (Cps2F product) and assay of Cps2G activity in the absence of other glycosyltransferases. Four independent reactions of membranes containing Cps2E, Cps2T and Cps2F and were incubated as described above to synthesize the Cps2F product. Extracted glycolipids were pooled, dried, then and suspended in 100 µl of 100 mM Tris-acetate, pH 7.6 containing 0.1% NP-40. This solution was used as the source of Und-P-P-[^{14}C]Glc-Rha-Rha-Rha in subsequent glycosyltransferase reactions. Membranes (3 µg total protein) from DJ011 containing non-His-tagged Cps2G or from RC124, a vector-only *E. coli* strain, were incubated in reaction mixtures containing 0.008% NP-40, 10 mM $MnCl_2$, 1 mM DTT, 25 µl of the Und-P-P-[^{14}C]Glc-Rha-Rha-Rha solution, and 0.1 mM dTDP-Rha. Reaction products were visualized as described above.

Bioinformatic analysis of the Cps2E homologues in *S. pneumoniae*. Amino acid sequences from 94 serotypes (90 listed in reference (4), and the remaining 4 in references 6 and 7 were analyzed for Cps2E homologues. 69 serotypes contained Cps2E homologues and these amino acid sequences were aligned by ClustalW (Geneious version 5.6.4) and a consensus amino acid sequence was generated.

RESULTS

Functional analysis 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 (8, 17). We next sought to demonstrate enzyme activities of the remaining glycosyltransferases (Cps2F, Cps2G, Cps2I) and thus clarify the functions of the gene products whose deletions in selected for secondary mutations in *cps2E* (17). In *in vitro* glycosyltransferase assays, we utilized individual membrane preparations obtained from *E. coli* expressing *S. pneumoniae* glycosyltransferases Cps2F, Cps2G, or Cps2I. As previously reported (17), each of these glycosyltransferases is membrane-associated when expressed in *E. coli.* As for Cps2T (17), *in vitro* enzyme activity required the addition of non-ionic detergent (0.008% NP-40) at a concentration well below the critical micelle concentration. To identify the enzyme that catalyzes the third step in CPS synthesis, i.e. the second rhamnosyltransferase, we combined Cps2E- and Cps2T-containing membranes in reactions with membranes containing the indicated glycosyltransferases and nucleotide sugars. 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-^{14}C-Glc and dTDP-Rha, only the reaction that contained Cps2F resulted in the formation of a new ^{14}C-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*).

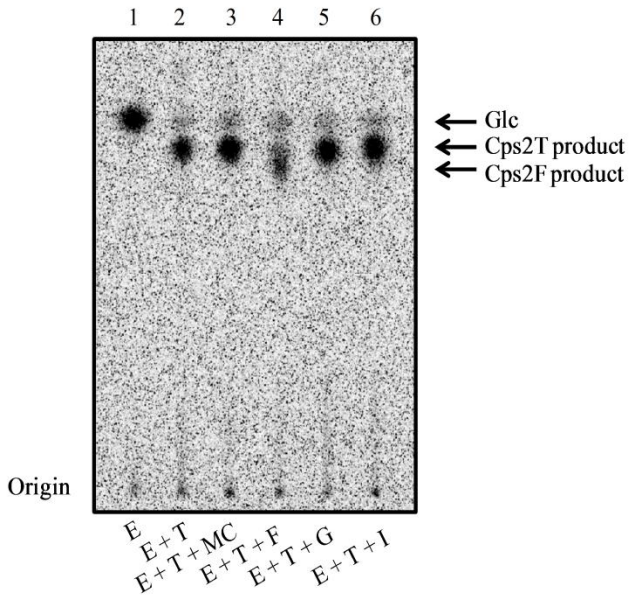

FIG 2. Cps2F rhamnosyltransferase activity. Cps2E-containing membranes (3 μg total protein) were incubated with membranes containing the indicated non His-tagged glycosyltransferases (3 μg total protein) or membrane controls (MC; 3 μg total protein from vector only *E. coli* strain RC124) and UDP-^{14}C-Glc only or UDP-^{14}C-Glc and dTDP-Rha. Reactions (75 μl total volume) were incubated for 1 h at 10^{o}C and contained 0.008% NP-40, 10 mM $MnCl_2$, 1 mM DTT, 0.075 μCi 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 the Materials and Methods. E, Cps2E; T, Cps2T; F, Cps2F; G, Cps2G; I, Cps2I. Reaction(s): 1, UDP-^{14}C-Glc only; 2-6, UDP-^{14}C-Glc and dTDP-Rha.

The third and fourth sugars, and their glycosidic linkages, are identical in the type 2 repeat unit (18). Cps2F may then catalyze the addition of two consecutive α1-3 Rha residues as predicted by previous analyses (1, 17). Separation of the Cps2T and Cps2F dependent products by the indicated TLC method, however, was insufficient to ascertain whether Cps2F synthesized two identifiable products. To improve TLC resolution, several parameters were adjusted including the polarity of the mobile phase, length of run, silica matrix composition and separation in 2-dimentions. None of these modifications greatly increased the separation of the Cps2F and Cps2T products (*data not*

shown). Data demonstrating Cps2G activity, described below, provides evidence for a Cps2F dependent addition of two consecutive α1-3 Rha residues.

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. Here, only reactions with Cps2G resulted in a new ^{14}C-Glc labeled product (Fig. 3A), indicating that Cps2G catalyzes the fourth step in CPS synthesis. Since both UDP-^{14}C-Glc and dTDP-Rha were added to the reaction, we next sought to identify the sugar specificity of Cps2G. Here, reactions 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, and not Rha, to the Cps2F product indicating that Cps2G is the α1-2 glucosyltransferase. These results also provide evidence for the identity of the Cps2F reaction product to be Und-P-P-Glc-Rha-Rha-Rha.

Using the above approach, we attempted to assay enzyme activity for Cps2I. Several attempts, however, resulted in no new ^{14}C-Glc labeled product (Fig. 4A). We also attempted to assay Cps2I activity directly by measuring the incorporation of ^{3}H-GlcUA, from UDP-^{3}H-GlcUA, to the Cps2G reaction product. Here, we combined Cps2E-, Cps2T-, Cps2F-, Cps2G- , and Cps2I containing membranes with UDP-Glc, dTDP-Rha and UDP-^{3}H-GlcUA. Glycolipids formed in these reactions were extracted, dried and analyzed for the incorporation of ^{3}H-GlcUA.

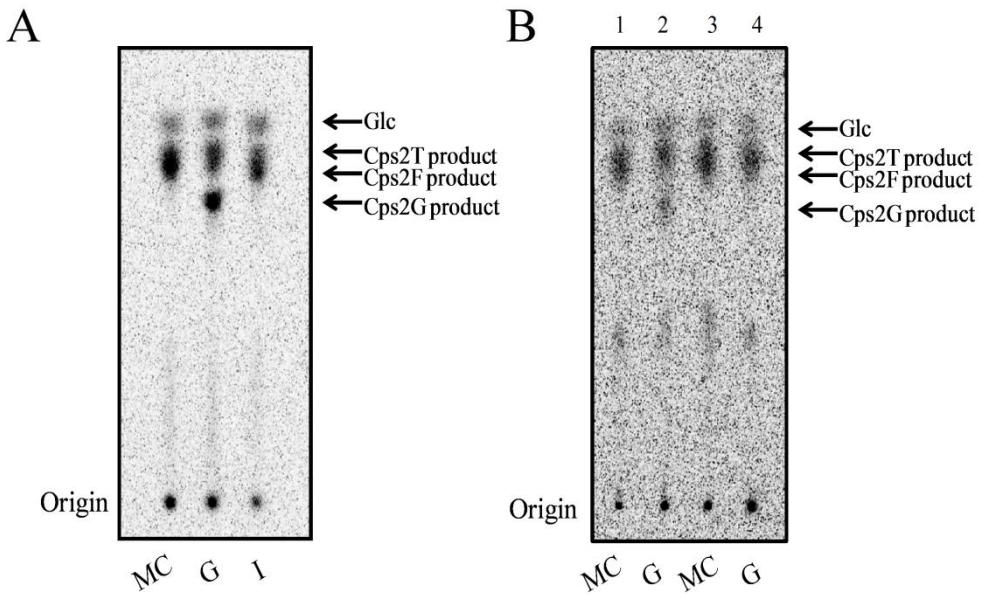

FIG 3. Cps2G glucosyltransferase activity. (A) Cps2E-, Cps2T-, and Cps2F-containing membranes (3 μg total protein each) were incubated with UDP-^{14}C-Glc, dTDP-Rha and membranes containing the non His-tagged Cps2G, Cps2I or membrane controls (MC; 3 μg total protein from vector only *E. coli* strain RC124). Reactions were processed as described in Fig 1. (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 non His-tagged Cps2G or membrane controls in the presence of UDP-Glc (reactions 1 and 2) or dTDP-Rha (reactions 3 and 4). Reactions contained 3 μg of total membrane protein, 0.008% NP-40, 10 mM $MnCl_2$, 1 mM DTT, 25 μl 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. G, Cps2G; I, Cps2I.

Membranes containing Cps2I did not incorporate ^{3}H-GlcUA above negative controls (identical reactions replacing Cps2I membranes with control membranes), in contrast to the positive control (*S. pneumoniae* WU2 membranes, containing the synthase Cps3S (9, 11), utilizing UDP-^{3}H-GlcUA) which incorporated ^{3}H-GlcUA (*data not shown*). Similar to our result, others have reported an inability to demonstrate terminal sugar addition for polysaccharide synthesis in various organisms (12, 21, 32). With one enzyme and one linkage unaccounted for, we predict that Cps2I adds the terminal α1-6 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 (1), and α1-6 GlcUA is a retaining linkage. The inability to demonstrate Cps2I activity may suggest that an addition factor is required for Cps2I activity *in vitro* and is considered in more detail in the Discussion.

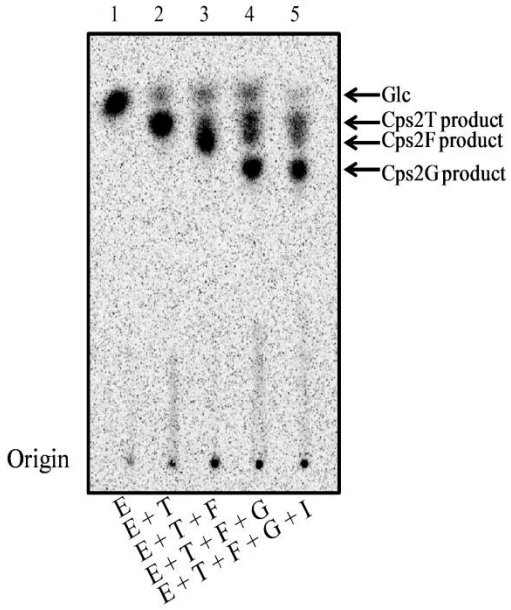

FIG 4. *S. pneumoniae* capsular glycosyltransferase activity.(A) Cps2E-containing membranes (3 μg total protein) were incubated with membranes containing the indicated non His-tagged glycosyltransferases (3 μg total protein) and nucleotide sugar. Reactions were processed as described in Fig 1 and contained 0.008% NP-40, 10 mM $MnCl_2$, 1 mM DTT, 0.075 μCi UDP-^{14}C-Glc (300 mCi/mmol) and, where indicated, 0.1 mM dTDP-Rha and/or 0.1 mM UDP-GlcUA. E, Cps2E; T, Cps2T; F, Cps2F; G, Cps2G; I, Cps2I. Reaction(s): 1, UDP-^{14}C-Glc only; 2-4, UDP-^{14}C-Glc and dTDP-Rha; 6, UDP-^{14}C-Glc, dTDP-Rha and UDP-GlcUA.

Amino acids selected for suppressor mutations in Cps2E correlate with highly conserved residues across CpsE homologues in *S. pneumoniae*. Deletion of the genes encoding glycosyltransferases (Cps2F, Cps2G, or Cps2I), the UDP-glucose dehydrogenase (Cps2K), the Wzx translocase (Cps2J), or the Wzy polymerase (Cps2H)

are lethal in the absence of secondary mutations (17, 42). Most of these secondary mutations map to *cps2E* and were primarily located in the protein cytoplasmic domain, but also in the extracytoplasmic loop domain (17, 42) (missense mutations shown in Fig.5A).

Cps2E homologues are present in at least 69 *S. pneumoniae* serotypes. To evaluate whether residues selected as suppressor mutations in Cps2E were also present 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 21 missense suppressor mutations in Cps2E and 17 of these are in residues that are 100% conserved in the 69 serotypes (Fig. 5B). Using the Fisher's Exact Test, a significant correlation ($P = 0.0031$) was observed between amino acids that are 100% conserved among *S. pneumoniae* Cps2E homologues and those selected as suppressor mutations. This result indicates that residues selected for suppressor mutations in Cps2E are not random, but identify important and highly conserved functional residues.

Suppressor mutations in the extracytoplasmic loop reduce Cps2E activity *in vitro*. Suppressor mutations mapping to the C terminal domain are likely to have a direct inhibition on Cps2E Glc-1-P transferase activity, as this domain alone contains glycosyltransferase activity in other members of the PHPT family (26, 27, 30, 37). However, suppressor mutations that map to the extracytoplasmic loop are more difficult to interpret but must have a function in limiting the accumulation of incomplete repeat units.

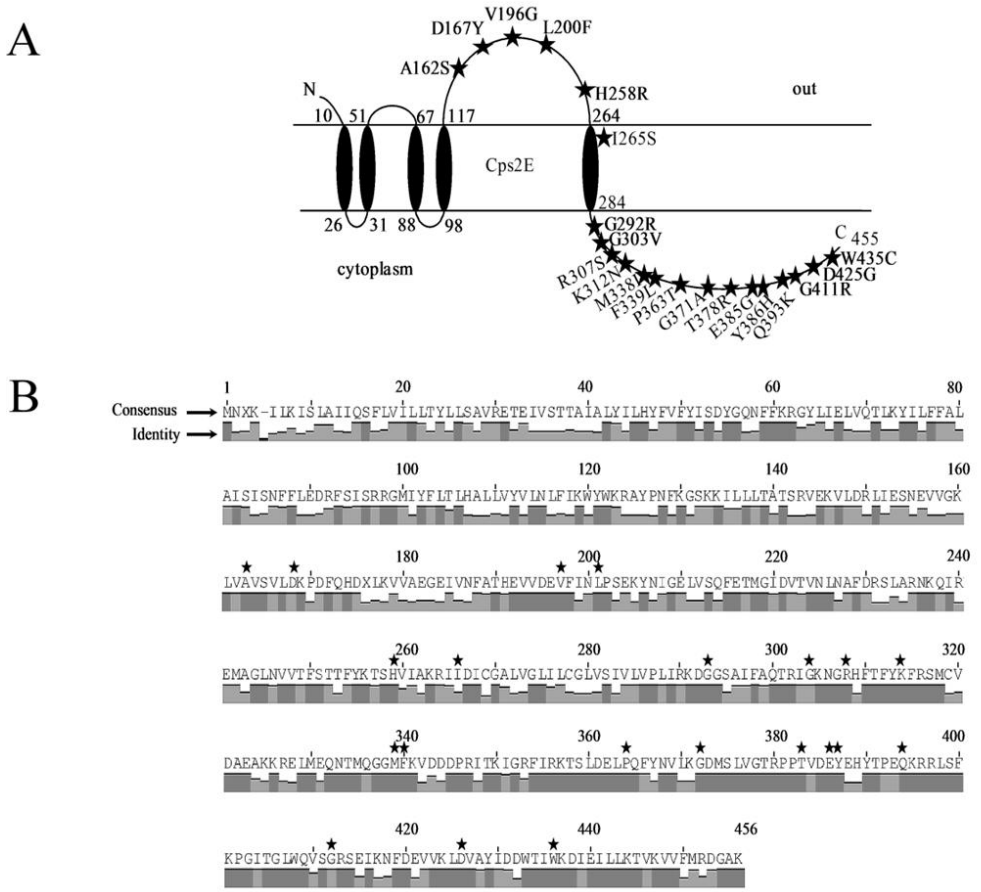

FIG 5. *S. pneumoniae* Cps2E topology and CpsE consensus sequence. (A) *S. pneumoniae* Cps2E topology with indicated missense suppressor mutations. (B) Amino acid sequences from 69 *S. pneumoniae* serotypes, each containing Cps2E homologues, were used in a ClustalW alignment to generate a consensus sequence. Bars indicate the degree of conservation for the respective residues, where the height of the bar correlates to a greater conservation (0 – 100%; and dark gray bars are 100% conserved). Stars above amino acids indicate residues in Cps2E where suppressor mutations were identified.

Five strains (DJ916 [Δ*cps2F*], DJ921 [Δ*cps2G*], BX547 [Δ*cps2K*], BX552 [Δ*cps2H*], BX605 [Δ*cps2K*]) from our collection each contain different suppressor mutations that map to the Cps2E extracytoplasmic loop (Table 1 and Fig. 5A). Δ*cps2K* mutants (BX547, BX605) in addition to the Δ*cps2G* (DJ921) mutant synthesize reduced levels of capsule that are present only in the membrane fraction (17, 42) indicating that

mutations in the extracytoplasmic loop maintain some ability to initiate capsule synthesis *in vivo*. We sought to utilize these strains to evaluate how the extracytoplasmic loop mutations affect Cps2E abundance and activity.

We evaluated the relative abundance of Cps2E in the membrane fraction of each of these mutants and found that protein levels vary according to the extracytoplasmic loop mutation (Fig. 6A). These membrane preparations were utilized in *in vitro* glycosyltransferase assays to determine Cps2E activity. Here *S. pneumoniae* membranes normalized to total membrane protein were incubated with UDP-[^{3}H]-Glc and the incorporation of [^{3}H]-Glc to the membrane acceptor was measured by liquid scintillation counting. Radioactive counts per minute (CPM) were normalized to the densitometry value of Cps2E protein levels and presented as relative Cps2E units (REU). Like mutations in the cytoplasmic domain (42), all Cps2E extracytoplasmic loop mutants had greatly reduced activities relative to parental Cps2E (Fig. 6B). Enzyme activity of the extracytoplasmic loop mutants were significantly different from that of negative and positive controls. No significant difference was observed when Cps2E extracytoplasmic loop mutants were expressed and assayed in an *E. coli* background (*data not shown*), suggesting that these environments may differently influence the activity of the extracytoplasmic loop mutants.

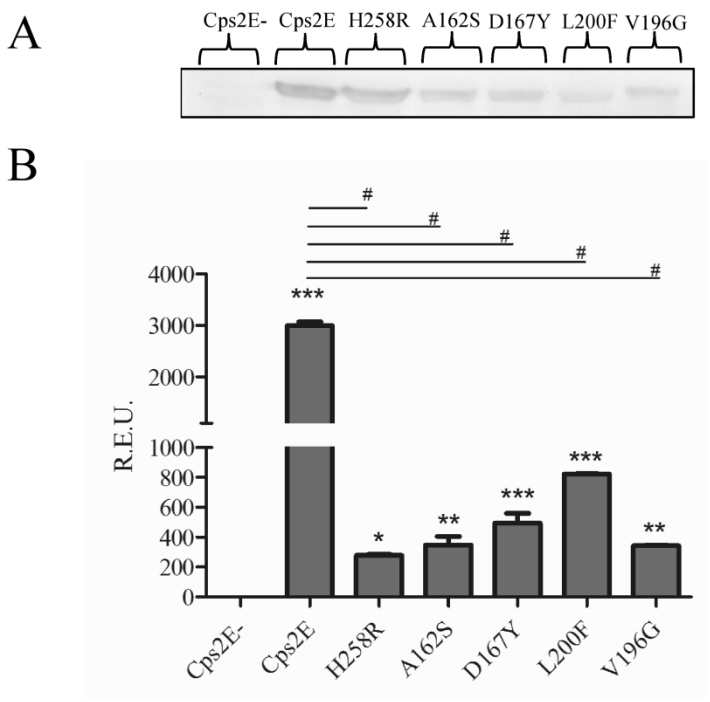

FIG 6. Cps2E protein levels and glycosyltransferase activity. (A) Cps2E abundance in membranes isolated from *S. pneumoniae* strains. Cps2E- (KA1522), Cps2E+ (D39), H258R (DJ916), A162S (DJ921), D167Y (BX547), L200F (BX552), V196G (BX605). 5 μg total membrane protein were separated by SDS-12% PAGE, transferred to a nitrocellulose membrane, and probed with anti-Cps2E antibody. (B) Cps2E activity in *S. pneumoniae* membranes. *S. pneumoniae* membranes (0.1 μg total protein) were used to measure incorporation of [^{3}H]-Glc from UDP-[^{3}H]-Glc to an organically soluble product. Reactions (75 μl total volume) were incubated for 1 h at 10^{o}C and contained 10 mM $MnCl_2$, 1 mM DTT, 0.075 μCi UDP-^{3}H-Glc (300 mCi/mmol). Glycolipids were extracted, dried and incorporated [^{3}H]Glc was analyzed by liquid scintillation counting. CPM; counts per min. The results are the means (plus standard errors) of duplicate samples and representative of at least two independent membrane preparations. One way Analysis of Variance (ANOVA) was used to calculate *P* values. *, $P < 0.05$; **, $P < 0.005$; *** $P < 0.001$ when compared to Cps2E-. #, $P < 0.001$ when compared to Cps2E+. Radioactive counts per minute (CPM) were normalized to the densitometry value of Cps2E protein levels and presented as relative Cps2E units (REU).

DISCUSSION

The *S. pneumoniae* type 2 capsule repeat unit is a hexasaccharide structure that is synthesized by the activity of multiple glycosyltransferases. Cps2E and Cps2T catalyze the first two linkages. In addition to Cps2E and Cps2T, the *cps2* locus encodes three additional glycosyltransferases (Cps2F, Cps2G, and Cps2I) that are responsible for catalyzing the remaining four glycosidic linkages to complete the capsule repeat unit. This observation, that three enzymes catalyze four linkages, suggested that a single glycosyltransferase encoded within the *cps* locus is capable of adding two sugar residues. Aanensen *et al.* identified that *cps2F* (also annotated as *wchG*) is the only inverting transferase gene in the *cps* locus of serotype 2, which has two identical inverting linkages in the CPS structure: Rhaα1-3Rhaα1-3 (1). Our results demonstrating the activities of Cps2F and Cps2G provide evidence that Cps2F is responsible for the sequential addition of two α1-3 linked Rha residues. Troutman and Imperiali provided the first mechanistic characterization of a bacterial glycosyltransferase, *Salmonella enterica* PglH, that sequentially adds multiple sugars: three α1-4 N-acetylgalactosamine residues (33). A processive and dissociative mechanism was proposed which served as a "counting mechanism" to limit the amount of sugars added. A nine amino acid motif (EX_7E, any seven amino acids flanked by two glutamic acid residues) was found to be important in PglH ability to sequentially add multiple residues (33). We predict that Cps2F, also containing an EX_7E motif, adopts a similar mechanism to sequentially add two α1-3Rha residues. Additional studies, however, will be required to address the Cps2F mechanism of sugar transfer.

We were unable to experimentally demonstrate that Cps2I performs the final GlcUA addition to the growing repeat unit. Others have reported a similar finding when biochemically characterizing the addition of terminal sugars on membrane linked polysaccharides from *E. coli*, *Klebsiella aerogenes*, and *S. pneumoniae* serotype 14 (12, 21, 32). In each of these studies, utilization of membrane associated glycosyltransferases in *in vitro* assays could not detect addition the terminal sugar (GlcUA in *E. coli* and *K. aerogenes*, and galactose in *S. pneumoniae*); a result that is speculated to be a limitation of the *in vitro* system. In the case of Cps2I, a defined protein modification/substrate/chaperone may be required to demonstrate activity *in vitro*. Addition of cell lysates containing functional Cps2K (UDP-Glc dehydrogenase) did not enhance detection of a Cps2I dependent sugar addition (*data not shown*). An analysis of the *S. pneumoniae* phosphoproteome has identified a phosphopeptide that corresponds to Cps2I (31). It remains possible then that *in vivo* Cps2I is phosphorylated by a currently unknown kinase which allows full functionality.

The initiating capsular glycosyltransferase Cps2E belongs to the phosphate hexose-1-phosphate transferase (PHPT) family of enzymes. Much of our understanding of PHPT enzymes comes from studies of *S. enterica* WbaP, the initiating O-antigen glycosyltransferase and prototypical PHPT enzyme. WbaP possess three domains the N terminal- , central- (extracytoplasmic loop), and C terminal-domain (30, 37, 38). Only the C terminal domain is required for sugar transfer in WbaP (26, 30, 37). Deletion of the extracytoplasmic loop domain in WbaP maintained repeat unit initiation, but resulted in an altered distribution of the chain-length (30). This suggested that PHPT enzymes are involved in more than just initiation of repeat unit synthesis, but may play a role in

additional modulator functions. The physical location of the extracytoplasmic loop also suggests a function that is necessary for either interaction with other surface enzymes or with an environmental factor.

In our previous studies (17, 42), we identified 5 amino acid residues (A162, D167, V196, L200, H258) in the extracytoplasmic loop of *S. pneumoniae* Cps2E that were selected as suppressor mutations (A162S, D167Y, V196G, L200F, H258R). Each of these mutations results in a non conservative amino acid change affecting either polarity or presence of an aromatic side chain. The extracytoplasmic loops of *S. enterica* WbaP and Cps2E are both 148 amino acids in length but share limited sequence identity; identities = 23/152 (15%), positives = 49/152 (32%), gaps = 8/152 (5%). Despite this, two amino acid residues selected as suppressor mutations in Cps2E (A162, D167) are also conserved in their relative position and identity in WbaP (A177, D182) *(data not shown).*The importance of these residues in WbaP, however, remains to be evaluated. In *in vitro* glycosyltransferase assays using *S. pneumoniae* membranes, we confirmed that mutation of these extracytoplasmic loop residues significantly reduced, but did not completely abolish, activity. This is in contrast to expression and analysis of Cps2E extracytoplasmic loop mutants in an *E. coli* background, where mutants were not significantly different from background controls (*data not shown*). Ongoing experiments will address the direct effect of the suppressor mutations in the extracytoplasmic loop of Cps2E on known factors (environmental and enzymatic) that affect capsule levels and distribution.

REFERENCES

1. **Aanensen, D. M., A. Mavroidi, S. D. Bentley, P. R. Reeves, and B. G. Spratt.** 2007. Predicted functions and linkage specificities of the products of the *Streptococcus pneumoniae* capsular biosynthetic loci. J Bacteriol **189:**7856-76.
2. **Avery, O. T., and R. Dubos.** 1931. The protective action of a specific enzyme against type III pneumococcus infection in mice. J Exp Med **54:**73-89.
3. **Avery, O. T., C. M. MacLeod, and M. McCarty.** 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:**137-58.
4. **Bentley, S. D., D. M. Aanensen, A. Mavroidi, D. Saunders, E. Rabbinowitsch, M. Collins, K. Donohoe, D. Harris, L. Murphy, M. A. Quail, G. Samuel, I. C. Skovsted, M. S. Kaltoft, B. Barrell, P. R. Reeves, J. Parkhill, and B. G. Spratt.** 2006. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. PLoS Genet **2:**e31.
5. **Brown, E. J., K. A. Joiner, R. M. Cole, and M. Berger.** 1983. Localization of complement component 3 on *Streptococcus pneumoniae*: anti-capsular antibody causes complement deposition on the pneumococcal capsule. Infect Immun **39:**403-9.
6. **Calix, J. J., and M. H. Nahm.** 2010. A new pneumococcal serotype, 11E, has a variably inactivated *wcjE* gene. J Infect Dis **202:**29-38.
7. **Calix, J. J., R. J. Porambo, A. M. Brady, T. R. Larson, J. Yother, C. Abeygunwardana, and M. H. Nahm.** 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:**27885-94.
8. **Cartee, R. T., W. T. Forsee, M. H. Bender, K. D. Ambrose, and J. Yother.** 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:**7425-33.
9. **Cartee, R. T., W. T. Forsee, J. S. Schutzbach, and J. Yother.** 2000. Mechanism of type 3 capsular polysaccharide synthesis in *Streptococcus pneumoniae*. J Biol Chem **275:**3907-14.
10. **Dillard, J. P., M. W. Vandersea, and J. Yother.** 1995. Characterization of the cassette containing genes for type 3 capsular polysaccharide biosynthesis in *Streptococcus pneumoniae*. J Exp Med **181:**973-83.
11. **Dillard, J. P., and J. Yother.** 1994. Genetic and molecular characterization of capsular polysaccharide biosynthesis in *Streptococcus pneumoniae* type 3. Mol Microbiol **12:**959-72.
12. **Drummelsmith, J., and C. Whitfield.** 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:**1321-32.
13. **Francis, T., and W. S. Tillett.** 1930. Cutaneous reactions in pneumonia. The development of antibodies following the intradermal injection of type-specific polysaccharide. J Exp Med **52:**573-85.

14. **Griffith, F.** 1928. The significance of Pneumococcal types. J Hyg (Lond) **27:**113-59.
15. **Hardy, G. G., A. D. Magee, C. L. Ventura, M. J. Caimano, and J. Yother.** 2001. Essential role for cellular phosphoglucomutase in virulence of type 3 *Streptococcus pneumoniae*. Infect Immun **69:**2309-17.
16. **Iannelli, F., B. J. Pearce, and G. Pozzi.** 1999. The type 2 capsule locus of *Streptococcus pneumoniae*. J Bacteriol **181:**2652-4.
17. **James, D. B., and J. Yother.** 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 beta1-4 Rhamnose, the second sugar residue in the repeat unit. J Bacteriol **194:**6479-89.
18. **Jansson, P. E., B. Lindberg, M. Anderson, U. Lindquist, and J. Henrichsen.** 1988. Structural studies of the capsular polysaccharide from *Streptococcus pneumoniae* type 2, a reinvestigation. Carbohydr Res **182:**111-7.
19. **Jiang, X. M., B. Neal, F. Santiago, S. J. Lee, L. K. Romana, and P. R. Reeves.** 1991. Structure and sequence of the rfb (O antigen) gene cluster of *Salmonella* serovar typhimurium (strain LT2). Mol Microbiol **5:**695-713.
20. **Kolkman, M. A., D. A. Morrison, B. A. Van Der Zeijst, and P. J. Nuijten.** 1996. The capsule polysaccharide synthesis locus of *Streptococcus pneumoniae* serotype 14: Identification of the glycosyl transferase gene *cps14E*. J Bacteriol **178:**3736-41.
21. **Kolkman, M. A., W. Wakarchuk, P. J. Nuijten, and B. A. van der Zeijst.** 1997. Capsular polysaccharide synthesis in *Streptococcus pneumoniae* serotype 14: molecular analysis of the complete *cps* locus and identification of genes encoding glycosyltransferases required for the biosynthesis of the tetrasaccharide subunit. Mol Microbiol **26:**197-208.
22. **Llull, D., E. Garcia, and R. Lopez.** 2001. Tts, a processive beta-glucosyltransferase of *Streptococcus pneumoniae*, directs the synthesis of the branched type 37 capsular polysaccharide in pneumococcus and other gram-positive species. J Biol Chem **276:**21053-61.
23. **Macleod, C. M., R. G. Hodges, M. Heidelberger, and W. G. Bernhard.** 1945. Prevention of pneumococcal pneumonia by immunization with specific capsular polysaccharides. J Exp Med **82:**445-65.
24. **Magee, A. D., and J. Yother.** 2001. Requirement for capsule in colonization by *Streptococcus pneumoniae*. Infect Immun **69:**3755-61.
25. **Nelson, A. L., A. M. Roche, J. M. Gould, K. Chim, A. J. Ratner, and J. N. Weiser.** 2007. Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. Infect Immun **75:**83-90.
26. **Patel, K. B., E. Ciepichal, E. Swiezewska, and M. A. Valvano.** 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:**116-22.
27. **Pelosi, L., M. Boumedienne, N. Saksouk, J. Geiselmann, and R. A. Geremia.** 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:**857-65.

28. **Centers for Disease Control and Prevention.**7 May 2012, revision date. Epidemiology and prevention of vaccine-preventable diseases, 12th ed, second printing, p 233-248. Public Health Foundation, Washington, DC. http://www.cdc.gov/vaccines/pubs/pinkbook/table-of-contents.html.
29. **Raetz, C. R., and C. Whitfield.** 2002. Lipopolysaccharide endotoxins. Annu Rev Biochem **71:**635-700.
30. **Saldias, M. S., K. Patel, C. L. Marolda, M. Bittner, I. Contreras, and M. A. Valvano.** 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:**440-53.
31. **Sun, X., F. Ge, C. L. Xiao, X. F. Yin, R. Ge, L. H. Zhang, and Q. Y. He.** 2010. Phosphoproteomic analysis reveals the multiple roles of phosphorylation in pathogenic bacterium *Streptococcus pneumoniae*. J Proteome Res **9:**275-82.
32. **Sutherland, I. W., and M. Norval.** 1970. The synthesis of exopolysaccharide by *Klebsiella aerogenes* membrane preparations and the involvement of lipid intermediates. Biochem J **120:**567-76.
33. **Troutman, J. M., and B. Imperiali.** 2009. *Campylobacter jejuni* PglH is a single active site processive polymerase that utilizes product inhibition to limit sequential glycosyl transfer reactions. Biochemistry **48:**2807-16.
34. **Valvano, M. A.** 2003. Export of O-specific lipopolysaccharide. Front Biosci **8:**s452-71.
35. **van Dam, J. E., A. Fleer, and H. Snippe.** 1990. Immunogenicity and immunochemistry of *Streptococcus pneumoniae* capsular polysaccharides. Antonie Van Leeuwenhoek **58:**1-47.
36. **van Selm, S., M. A. Kolkman, B. A. van der Zeijst, K. A. Zwaagstra, W. Gaastra, and J. P. van Putten.** 2002. Organization and characterization of the capsule biosynthesis locus of *Streptococcus pneumoniae* serotype 9V. Microbiology **148:**1747-55.
37. **Wang, L., D. Liu, and P. R. Reeves.** 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:**2598-604.
38. **Wang, L., and P. R. Reeves.** 1994. Involvement of the galactosyl-1-phosphate transferase encoded by the *Salmonella enterica* rfbP gene in O-antigen subunit processing. J Bacteriol **176:**4348-56.
39. **Weiser, J. N., R. Austrian, P. K. Sreenivasan, and H. R. Masure.** 1994. Phase variation in pneumococcal opacity: relationship between colonial morphology and nasopharyngeal colonization. Infect Immun **62:**2582-9.
40. **Whitfield, C.** 2006. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. Annu Rev Biochem **75:**39-68.
41. **Wood, W. B., Jr., and M. R. Smith.** 1949. The inhibition of surface phagocytosis by the capsular slime layer of pneumococcus type III. J Exp Med **90:**85-96.
42. **Xayarath, B., and J. Yother.** 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:**3369-81.

43. **Yother, J.** 2011. Capsules of *Streptococcus pneumoniae* and other bacteria: paradigms for polysaccharide biosynthesis and regulation. Annu Rev Microbiol **65:**563-81.

44. **Yuasa, R., M. Levinthal, and H. Nikaido.** 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:**433-44.

SUMMARY AND DISCUSSION

The polysaccharide capsule of *S. pneumoniae* is a bonafide virulence factor. Multiple studies across decades of research have revealed that this surface structure influences nasopharyngeal colonization and systemic virulence in animal models of pneumococcal disease (4, 30, 32, 54, 61). *S. pneumoniae* capsule research has resulted in the development of polysaccharide-based vaccines that have significantly reduced, but not eradicated, the burden of pneumococcal disease. As a result, ongoing efforts are aimed at increasing our understanding of the mechanisms of capsule synthesis and its regulation in order to develop more effective therapeutics. Therapeutic targets that focus on the conserved mechanisms of capsule synthesis, in contrast to the varying capsule types, may prove to be more effective in controlling pneumococcal disease.

In the Wzy-dependent pathway, multiple enzymes known as glycosyltransferases sequentially build the repeat unit on a membrane lipid acceptor. These glycosyltransferases utilize nucleotide sugars as donors to transfer a glycosyl moiety to the acceptor. In *S. pneumoniae* serotype 2 capsule synthesis, these glycosyltransferases are Cps2E, Cps2T, Cps2F, Cps2G, and Cps2I (19, 37, 38).

Much of our understanding of glycosyltransferases in the Wzy-dependent pathway comes from study of LPS O-antigen biosynthesis in Gram-negative bacteria, where the pathway was first discovered (70, 80). O-antigen, like the capsule repeat unit, is synthesized on Und-P by multiple glycosyltransferases. Early studies investigating the genetics of LPS biosynthesis identified that *S. enterica* mutants deficient in O-antigen side chain assembly do not polymerize O-antigen (64). These mutants also contained

secondary mutations. Subsequent studies have identified similar findings in other surface structures, including capsules and teichoic acids, from numerous bacterial species (13, 25, 43, 72, 98). In these reports, deletion of certain genes involved in the synthesis of the respective lipid linked polysaccharide repeat unit is lethal in the absence of a secondary mutation. By utilizing this conditional lethality, we have demonstrated that the committed step, i.e., the first physiologically irreversible step, in *S. pneumoniae* type 2 capsule synthesis is catalyzed by the second glycosyltransferase (Cps2T) (38). Our results suggest that this critical step is not limited to serotype 2 or *S. pneumoniae*, but possibly all Wzy-dependent glycosyltransferases that function immediately after the readily reversible initiating glycosyltransferase. A common principle of many biochemical pathways is that the committed step serves as an important point of control, and thus capable of exploitation for therapeutic interventions.

Another conserved property in the Wzy-pathway is the enzyme that initiates repeat unit synthesis. In *S. pneumoniae*, and many Gram-negative species that utilize the Wzy-pathway, polysaccharide repeat unit initiation begins with a phosphate hexose-1-phosphate transferase (PHPT). PHPT family members possess three domains the N terminal- , central- (extracytoplasmic loop), and C terminal-domain. Wang *et al.* were the first to show that the PHPT enzyme WbaP (*S. enterica*) is a bifunctional enzyme possessing two domains (87, 88). They proposed that the C-terminal domain is responsible for sugar transfer and the N-terminal region is required for release of the lipid-linked sugar after WbaP catalysis (87, 88). Later studies by Saldias *et al.* provided additional evidence that in WbaP, the C-terminal domain alone catalyzes sugar transfer (65, 74). They also identified an additional domain in WbaP, dividing the previously

described N-terminal region into a new domain: the extracytoplasmic loop (74). Deletion of the extracytoplasmic loop domain in WbaP maintained repeat unit initiation, but resulted in an altered distribution of the polysaccharide chain-length (74). This result suggested that PHPT enzymes are involved in more than just initiation of repeat unit synthesis, and may play a role in additional functions required to fully modulate repeat unit synthesis.

In previous studies, we observed in that secondary suppressor mutations in the extracytoplasmic loop of Cps2E could limit the lethal accumulation of incomplete repeat units caused by the primary mutations in *S. pneumoniae* (38, 96). These results suggested that this domain plays an important role in capsule initiation and possibly the regulation of initiation. Our findings demonstrate that amino acids selected for suppressor mutations in the extracytoplasmic loop of Cps2E (A162, D167, V196, L200, H258) are highly conserved among *S. pneumoniae* serotypes and two (A162, D167) are conserved in their relative positions in *S. enterica* WbaP (A177, D182). *S. pneumoniae* membranes expressing Cps2E extracytoplasmic loop mutants were severely attenuated in their ability to transfer ^{3}H-Glc to the lipid acceptor *in vitro*, indicating that these residues are critical for Cps2E glycosyltransferase activity. The physical locations of extracytoplasmic loop of Cps2E also suggests a function that is necessary for either interaction with other surface enzymes or with an environmental factor. Ongoing experiments will address the direct effect of the suppressor mutations in the extracytoplasmic loop of Cps2E on known factors (environmental and enzymatic) that affect capsule levels.

Multiple factors, environmental and enzymatic, appear to influence capsule levels in *S. pneumoniae*. Environmental factors include the concentration of metabolizable

carbohydrates/growth media (63) and extracellular oxygen/oxidants (92), while enzymatic regulators include CpsABCD (8, 9, 59, 92) and the transcriptional regulator CcpA (20, 44). As introduced earlier, capsule levels are dynamic and depend upon the site of infection. *In vivo*, it is possible that extracellular factors serve as the stimuli that trigger the corresponding intracellular enzymatic changes required to modulate capsule.

Potential points and mechanisms of capsular glycosyltransferase regulation in *S. pneumoniae* serotype 2

A common point of regulation among glycosyltransferases is to modulate the availability of nucleotide sugars (83). Glc concentrations directly influence Glc-6-P levels, which is a necessary precursor in the biosynthesis of many nucleotide sugars (see introduction). In *S. pneumoniae*, a positive correlation between Glc concentrations and capsule is evident upon analysis of the natural sites of infection. While colonizing the human nasopharynx, which normally lacks free Glc (67), *S. pneumoniae* displays reduced capsule (47, 54, 61, 91). This is in contrast to infection of systemic sites, such as the blood, where free Glc concentrations are high (67) and capsule is increased (4, 32, 52, 54). These observations could suggest that glucose concentrations influence capsule via modulating the nucleotide sugar concentrations and glycosyltransferase activity. Nucleotide sugar concentrations have been shown to affect glycosyltransferase activity and capsule levels in a *S. pneumoniae* strain (WU2, serotype 3) utilizing the synthase pathway (84), but this remains to be determined in any Wzy-dependent serotype.

Our results demonstrate that inhibition of enzymes beyond the committed step to capsule synthesis is lethal (38, 96). This result suggests that if a capsular

glycosyltransferase is regulated, then it must act early in repeat unit synthesis prior to the irreversible accumulation of incomplete glycolipids. In this regard, two possible points of regulation in *S. pneumoniae* type 2 include the activities of Cps2E, the initiating glycosyltransferase, and/or that of Cps2T, which catalyzes the second and committed step.

Regulating the concentrations of nucleotide sugars as a means to influence Cps2E activity appears unlikely as its sugar donor, UDP-Glc, is also used for other processes such as teichoic acid synthesis (76). Furthermore, capsule and teichoic acid levels appear to be inversely proportioned (47, 91). One possibility for regulating Cps2E activity lies at the transcriptional level, where disrupting expression of *cps2E* could ultimately affect capsule synthesis initiation. Support for this possibility is evident through findings in *Streptococcus agalactaie*, where expression of the initiating capsular glycosyltransferase CpsE may be regulated by the putative transcriptional regulator CpsA. Using electrophoretic mobility shift assays, Hanson *et. al.* demonstrated that CpsA directly binds to a promoter upstream of *cpsE* in addition to another promoter at the beginning of the *cps* locus (31). An insertion-deletion of *cpsA* leads to a decrease in capsule production in *S. agalactaie*. These results ultimately led the authors to conclude that CpsA acts as a positive regulator of capsule by direct interaction with capsule DNA to facilitate transcriptional changes. CpsA and CpsE in *S. agalactaie* are homologus to *S. pneumoniae* Cps2A and Cps2E (60) however no promoter upstream of *S. pneumoniae cps2E* has been experimentally identified (60).

Regulation of Cps2T activity by the concentration of its nucleotide sugar substrate (dTDP-Rha) is a possible mechanism for modulating capsule levels. As described in the

first chapter, the possibility of dTDP-Rha concentrations influencing Cps2T activity was gained from an interpretation of the transformation efficiency for acquiring glycosyltransferase mutants. In that report the Δ*cps2F* mutants, lacking the gene encoding the second rhamnosyltransferase, were acquired at a frequency similar to that of Δ*cps2T* mutants (38). We reasoned that in Δ*cps2F* mutants, accumulation of dTDP-Rha results in a less efficient Cps2T, a slower accumulation of incomplete repeat units and thus a transformation efficiency similar to that of Δ*cps2T* mutants. In *S. pneumoniae*, dTDP-Rha is utilized only in capsule synthesis and a correlation between the two is likely to exist.

A recent report provided evidence that bacteria do regulate dTDP-Rha biosynthesis, which ultimately affects the concentration of rhamnose-containing glycolipids. Here, Aguirre-Ramirez *et. al.* described in *Pseudomonas aeruginosa* that the quorum-sensing transcriptional regulator RhlR specifically activates transcription of the dTDP-Rha biosynthetic operon *rmlBDAC* during early stationary growth (3) when rhamnolipid levels are high (99). Using *lacZ* fusions to *rmlB*, they identified a RhlR specific sequence in the promoter region of *rmlBDAC* that was required for RhlR activation. RmlBDAC is homologous to Cps2LMNO, however no promoters have been identified upstream of *cps2LMNO* in *S. pneumoniae.*

This dissertation describes how *S. pneumoniae* serotype 2 synthesizes its capsular polysaccharide through the activities of five glycosyltransferases. By using genetic and biochemical analyses, we revealed some of the mechanistic intricacies required for proper synthesis. Inhibition of any enzyme beyond the committed step in capsule synthesis is lethal in the absence of secondary mutations. Mapping these secondary mutations has

identified critical and conserved residues in the initiating glycosyltransferase, which is part of a larger family of enzymes not present in humans. Continued analysis of the basic mechanisms of capsule synthesis should lead to better therapeutics and therapeutic targets against pneumococci and other significant bacterial pathogens.

GENERAL LIST OF REFERENCES

1. **Aanensen, D. M., A. Mavroidi, S. D. Bentley, P. R. Reeves, and B. G. Spratt.** 2007. Predicted functions and linkage specificities of the products of the *Streptococcus pneumoniae* capsular biosynthetic loci. J Bacteriol **189:**7856-76.
2. **Abeyta, M., G. G. Hardy, and J. Yother.** 2003. Genetic alteration of capsule type but not PspA type affects accessibility of surface-bound complement and surface antigens of *Streptococcus pneumoniae*. Infect Immun **71:**218-25.
3. **Aguirre-Ramirez, M., G. Medina, A. Gonzalez-Valdez, V. Grosso-Becerra, and G. Soberon-Chavez.** 2012. The *Pseudomonas aeruginosa rmlBDAC* operon, encoding dTDP-L-rhamnose biosynthetic enzymes, is regulated by the quorum-sensing transcriptional regulator RhlR and the alternative sigma factor sigmaS. Microbiology **158:**908-16.
4. **Avery, O. T., and R. Dubos.** 1931. The protective action of a specific enzyme against type III pneumococcus infection in mice. J Exp Med **54:**73-89.
5. **Avery, O. T., C. M. MacLeod, and M. McCarty.** 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:**137-58.
6. **Barnes, D. M., *et al.*** 1995. Transmission of multidrug-resistant serotype 23F *Streptococcus pneumoniae* in group day care: evidence suggesting capsular transformation of the resistant strain in vivo. J Infect Dis **171:**890-6.
7. **Barreteau, H., *et al.*** 2009. Quantitative high-performance liquid chromatography analysis of the pool levels of undecaprenyl phosphate and its derivatives in bacterial membranes. J Chromatogr B Analyt Technol Biomed Life Sci **877:**213-20.
8. **Bender, M. H., R. T. Cartee, and J. Yother.** 2003. Positive correlation between tyrosine phosphorylation of CpsD and capsular polysaccharide production in *Streptococcus pneumoniae*. J Bacteriol **185:**6057-66.
9. **Bender, M. H., and J. Yother.** 2001. CpsB is a modulator of capsule-associated tyrosine kinase activity in *Streptococcus pneumoniae*. J Biol Chem **276:**47966-74.
10. **Bentley, S. D., *et al.*** 2006. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. PLoS Genet **2:**e31.

11. **Breton, C., L. Snajdrova, C. Jeanneau, J. Koca, and A. Imberty.** 2006. Structures and mechanisms of glycosyltransferases. Glycobiology **16:**29R-37R.
12. **Brown, E. J., K. A. Joiner, R. M. Cole, and M. Berger.** 1983. Localization of complement component 3 on *Streptococcus pneumoniae*: anti-capsular antibody causes complement deposition on the pneumococcal capsule. Infect Immun **39:**403-9.
13. **Burrows, L. L., and J. S. Lam.** 1999. Effect of *wzx* (*rfbX*) mutations on A-band and B-band lipopolysaccharide biosynthesis in *Pseudomonas aeruginosa* O5. J Bacteriol **181:**973-80.
14. **Calix, J. J., and M. H. Nahm.** 2010. A new pneumococcal serotype, 11E, has a variably inactivated *wcjE* gene. J Infect Dis **202:**29-38.
15. **Caputto, R., L. F. Leloir, C. E. Cardini, and A. C. Paladini.** 1950. Isolation of the coenzyme of the galactose phosphate-glucose phosphate transformation. J Biol Chem **184:**333-50.
16. **Cardini, C. E., A. C. Paladini, R. Caputto, and L. F. Leloir.** 1950. Uridine diphosphate glucose: The coenzyme of the galactose-glucose phosphate isomerization. Nature **165:**191.
17. **Cartee, R. T., W. T. Forsee, M. H. Bender, K. D. Ambrose, and J. Yother.** 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:**7425-33.
18. **Cartee, R. T., W. T. Forsee, J. S. Schutzbach, and J. Yother.** 2000. Mechanism of type 3 capsular polysaccharide synthesis in *Streptococcus pneumoniae*. J Biol Chem **275:**3907-14.
19. **Cartee, R. T., W. T. Forsee, and J. Yother.** 2005. Initiation and synthesis of the *Streptococcus pneumoniae* type 3 capsule on a phosphatidylglycerol membrane anchor. J Bacteriol **187:**4470-9.
20. **Carvalho, S. M., T. G. Kloosterman, O. P. Kuipers, and A. R. Neves.** 2011. CcpA ensures optimal metabolic fitness of *Streptococcus pneumoniae*. PLoS One **6:**e26707.
21. **Center, K. J.** 2007. Prevenar vaccination: review of the global data, 2006. Vaccine **25:**3085-9.
22. **Coffey, T. J., *et al.*** 1991. Horizontal transfer of multiple penicillin-binding protein genes, and capsular biosynthetic genes, in natural populations of *Streptococcus pneumoniae*. Mol Microbiol **5:**2255-60.
23. **Coffey, T. J., *et al.*** 1998. Recombinational exchanges at the capsular polysaccharide biosynthetic locus lead to frequent serotype changes among natural isolates of *Streptococcus pneumoniae*. Mol Microbiol **27:**73-83.
24. **Croucher, N. J., *et al.*** 2011. Rapid pneumococcal evolution in response to clinical interventions. Science **331:**430-4.
25. **D'Elia, M. A., *et al.*** 2006. Lesions in teichoic acid biosynthesis in *Staphylococcus aureus* lead to a lethal gain of function in the otherwise dispensable pathway. J Bacteriol **188:**4183-9.
26. **DeAngelis, P. L.** 2002. Microbial glycosaminoglycan glycosyltransferases. Glycobiology **12:**9R-16R.

27. **Forsee, W. T., R. T. Cartee, and J. Yother.** 2000. Biosynthesis of type 3 capsular polysaccharide in *Streptococcus pneumoniae*. Enzymatic chain release by an abortive translocation process. J Biol Chem **275:**25972-8.
28. **Francis, T., and W. S. Tillett.** 1930. Cutaneous reactions in pneumonia. The development of antibodies following the intradermal injection of type-specific polysaccharide. J Exp Med **52:**573-85.
29. **Gram, H. C.** 1884. Über die isolierte färbung der schizomyceten in schnitt- und trockenpräparaten. Fortschritte der Medizin**:**185–89.
30. **Griffith, F.** 1928. The significance of Pneumococcal types. J Hyg (Lond) **27:**113-59.
31. **Hanson, B. R., *et al.*** 2012. Functional analysis of the CpsA protein of *Streptococcus agalactiae*. J Bacteriol **194:**1668-78.
32. **Hardy, G. G., A. D. Magee, C. L. Ventura, M. J. Caimano, and J. Yother.** 2001. Essential role for cellular phosphoglucomutase in virulence of type 3 *Streptococcus pneumoniae*. Infect Immun **69:**2309-17.
33. **Hausdorff, W. P., D. R. Feikin, and K. P. Klugman.** 2005. Epidemiological differences among pneumococcal serotypes. Lancet Infect Dis **5:**83-93.
34. **Heidelberger, M., and F. E. Kendall.** 1929. A quantitative study of the precipitin reaction between type III Pneumococcus polysaccharide and purified homologous antibody. J Exp Med **50:**809-23.
35. **Henrichsen, J.** 1995. Six newly recognized types of *Streptococcus pneumoniae*. J Clin Microbiol **33:**2759-62.
36. **Hyams, C., E. Camberlein, J. M. Cohen, K. Bax, and J. S. Brown.** 2010. The *Streptococcus pneumoniae* capsule inhibits complement activity and neutrophil phagocytosis by multiple mechanisms. Infect Immun **78:**704-15.
37. **Iannelli, F., B. J. Pearce, and G. Pozzi.** 1999. The type 2 capsule locus of *Streptococcus pneumoniae*. J Bacteriol **181:**2652-4.
38. **James, D. B., and J. Yother.** 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 beta1-4 Rhamnose, the second sugar residue in the repeat unit. J Bacteriol **194:**6479-89.
39. **Jansson, P. E., B. Lindberg, M. Anderson, U. Lindquist, and J. Henrichsen.** 1988. Structural studies of the capsular polysaccharide from *Streptococcus pneumoniae* type 2, a reinvestigation. Carbohydr Res **182:**111-7.
40. **Jedrzejas, M. J.** 2001. Pneumococcal virulence factors: structure and function. Microbiol Mol Biol Rev **65:**187-207; first page, table of contents.
41. **Jiang, X. M., *et al.*** 1991. Structure and sequence of the rfb (O antigen) gene cluster of *Salmonella* serovar typhimurium (strain LT2). Mol Microbiol **5:**695-713.
42. **Kadioglu, A., J. N. Weiser, J. C. Paton, and P. W. Andrew.** 2008. The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. Nat Rev Microbiol **6:**288-301.
43. **Katzen, F., *et al.*** 1998. *Xanthomonas campestris* pv. campestris gum mutants: effects on xanthan biosynthesis and plant virulence. J Bacteriol **180:**1607-17.

44. **Kaufman, G. E., and J. Yother.** 2007. CcpA-dependent and -independent control of beta-galactosidase expression in *Streptococcus pneumoniae* occurs via regulation of an upstream phosphotransferase system-encoding operon. J Bacteriol **189:**5183-92.
45. **Kelly, T., J. P. Dillard, and J. Yother.** 1994. Effect of genetic switching of capsular type on virulence of *Streptococcus pneumoniae*. Infect Immun **62:**1813-9.
46. **Kilian, M., J. Mestecky, and R. E. Schrohenloher.** 1979. Pathogenic species of the genus *Haemophilus* and *Streptococcus pneumoniae* produce immunoglobulin A1 protease. Infect Immun **26:**143-9.
47. **Kim, J. O., and J. N. Weiser.** 1998. Association of intrastrain phase variation in quantity of capsular polysaccharide and teichoic acid with the virulence of *Streptococcus pneumoniae*. J Infect Dis **177:**368-77.
48. **Kolkman, M. A., D. A. Morrison, B. A. Van Der Zeijst, and P. J. Nuijten.** 1996. The capsule polysaccharide synthesis locus of *Streptococcus pneumoniae* serotype 14: Identification of the glycosyl transferase gene *cps14E*. J Bacteriol **178:**3736-41.
49. **Kolkman, M. A., B. A. van der Zeijst, and P. J. Nuijten.** 1997. Functional analysis of glycosyltransferases encoded by the capsular polysaccharide biosynthesis locus of *Streptococcus pneumoniae* serotype 14. J Biol Chem **272:**19502-8.
50. **Kolkman, M. A., W. Wakarchuk, P. J. Nuijten, and B. A. van der Zeijst.** 1997. Capsular polysaccharide synthesis in *Streptococcus pneumoniae* serotype 14: molecular analysis of the complete *cps* locus and identification of genes encoding glycosyltransferases required for the biosynthesis of the tetrasaccharide subunit. Mol Microbiol **26:**197-208.
51. **Lairson, L. L., B. Henrissat, G. J. Davies, and S. G. Withers.** 2008. Glycosyltransferases: structures, functions, and mechanisms. Annu Rev Biochem **77:**521-55.
52. **MacLeod, C., and M. R. Kraus.** 1950. Relation of virulence of pneumococcal strains for mice to the quantity of capsular polysaccharide formed in vitro. J Exp Med **92:**1-9.
53. **Macleod, C. M., R. G. Hodges, M. Heidelberger, and W. G. Bernhard.** 1945. Prevention of pneumococcal pneumonia by immunization with specific capsular polysaccharides. J Exp Med **82:**445-65.
54. **Magee, A. D., and J. Yother.** 2001. Requirement for capsule in colonization by *Streptococcus pneumoniae*. Infect Immun **69:**3755-61.
55. **Matthias, K. A., A. M. Roche, A. J. Standish, M. Shchepetov, and J. N. Weiser.** 2008. Neutrophil-toxin interactions promote antigen delivery and mucosal clearance of Streptococcus pneumoniae. J Immunol **180:**6246-54.
56. **Mengin-Lecreulx, D., L. Texier, M. Rousseau, and J. van Heijenoort.** 1991. The *murG* gene of *Escherichia coli* codes for the UDP-N-acetylglucosamine: N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase involved in the membrane steps of peptidoglycan synthesis. J Bacteriol **173:**4625-36.

57. **Morona, J. K., D. C. Miller, R. Morona, and J. C. Paton.** 2004. The effect that mutations in the conserved capsular polysaccharide biosynthesis genes *cpsA*, *cpsB*, and *cpsD* have on virulence of *Streptococcus pneumoniae*. J Infect Dis **189:**1905-13.
58. **Morona, J. K., R. Morona, and J. C. Paton.** 1997. Characterization of the locus encoding the *Streptococcus pneumoniae* type 19F capsular polysaccharide biosynthetic pathway. Mol Microbiol **23:**751-63.
59. **Morona, J. K., J. C. Paton, D. C. Miller, and R. Morona.** 2000. Tyrosine phosphorylation of CpsD negatively regulates capsular polysaccharide biosynthesis in *Streptococcus pneumoniae*. Mol Microbiol **35:**1431-42.
60. **Moscoso, M., and E. Garcia.** 2009. Transcriptional regulation of the capsular polysaccharide biosynthesis locus of *Streptococcus pneumoniae*: a bioinformatic analysis. DNA Res **16:**177-86.
61. **Nelson, A. L., *et al.*** 2007. Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. Infect Immun **75:**83-90.
62. **Nesin, M., M. Ramirez, and A. Tomasz.** 1998. Capsular transformation of a multidrug-resistant *Streptococcus pneumoniae* in vivo. J Infect Dis **177:**707-13.
63. **Ogunniyi, A. D., P. Giammarinaro, and J. C. Paton.** 2002. The genes encoding virulence-associated proteins and the capsule of *Streptococcus pneumoniae* are upregulated and differentially expressed *in vivo*. Microbiology **148:**2045-53.
64. **Pasteur, L.** 1881. Note sur la maladie nouvelle provoquee par la salive d'un enfant mort de la rage. Bull. Acad, Med. (Paris) **10:**94-103.
65. **Patel, K. B., E. Ciepichal, E. Swiezewska, and M. A. Valvano.** 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:**116-22.
66. **Pelosi, L., M. Boumedienne, N. Saksouk, J. Geiselmann, and R. A. Geremia.** 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:**857-65.
67. **Philips, B. J., J. X. Meguer, J. Redman, and E. H. Baker.** 2003. Factors determining the appearance of glucose in upper and lower respiratory tract secretions. Intensive Care Med **29:**2204-10.
68. **Postma, P. W., J. W. Lengeler, and G. R. Jacobson.** 1993. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. Microbiol Rev **57:**543-94.
69. **Centers for Disease Control and Prevention.**7 May 2012, revision date. Epidemiology and prevention of vaccine-preventable diseases, 12th ed, second printing, p 233-248. Public Health Foundation, Washington, DC. http://www.cdc.gov/vaccines/pubs/pinkbook/table-of-contents.html.
70. **Raetz, C. R., and C. Whitfield.** 2002. Lipopolysaccharide endotoxins. Annu Rev Biochem **71:**635-700.
71. **Rao, S. T., and M. G. Rossmann.** 1973. Comparison of super-secondary structures in proteins. J Mol Biol **76:**241-56.
72. **Rick, P. D., *et al.*** 2003. Evidence that the *wzxE* gene of *Escherichia coli* K-12 encodes a protein involved in the transbilayer movement of a trisaccharide-lipid

intermediate in the assembly of enterobacterial common antigen. J Biol Chem **278:**16534-42.

73. **Saksouk, N., *et al.*** 2005. The capsular polysaccharide biosynthesis of *Streptococcus pneumoniae* serotype 8: functional identification of the glycosyltransferase WciS (Cap8H). Biochem J **389:**63-72.
74. **Saldias, M. S., *et al.*** 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:**440-53.
75. **Schiemann, O., and W. Casper.** 1927. Sind die spezifisch präcipitabelen Substanzen der 3 Pneumokokkentypen Haptene. Z Hyg Infektionskr **108:**220-57.
76. **Seo, H. S., R. T. Cartee, D. G. Pritchard, and M. H. Nahm.** 2008. A new model of pneumococcal lipoteichoic acid structure resolves biochemical, biosynthetic, and serologic inconsistencies of the current model. J Bacteriol **190:**2379-87.
77. **Sternberg, G. M.** 1881. A fatal form ofsepticaemia in the rabbit, produced by the subcutaneous injection of human saliva. Natl. Board Health Bull. **2:**781-783.
78. **Tomasz, A.** 1997. Antibiotic resistance in *Streptococcus pneumoniae*. Clin Infect Dis **24 Suppl 1:**S85-8.
79. **Tomasz, A., and R. D. Hotchkiss.** 1964. Regulation of the transformability of pneumococcal cultures by macromolecular cell products. Proc Natl Acad Sci U S A **51:**480-7.
80. **Valvano, M. A.** 2003. Export of O-specific lipopolysaccharide. Front Biosci **8:**s452-71.
81. **van Dam, J. E., A. Fleer, and H. Snippe.** 1990. Immunogenicity and immunochemistry of *Streptococcus pneumoniae* capsular polysaccharides. Antonie Van Leeuwenhoek **58:**1-47.
82. **van Selm, S., *et al.*** 2002. Organization and characterization of the capsule biosynthesis locus of *Streptococcus pneumoniae* serotype 9V. Microbiology **148:**1747-55.
83. **Varki, A.** 2009. Essentials of glycobiology, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
84. **Ventura, C. L., R. T. Cartee, W. T. Forsee, and J. Yother.** 2006. Control of capsular polysaccharide chain length by UDP-sugar substrate concentrations in *Streptococcus pneumoniae*. Mol Microbiol **61:**723-33.
85. **Walport, M. J.** 2001. Complement. First of two parts. N Engl J Med **344:**1058-66.
86. **Walport, M. J.** 2001. Complement. Second of two parts. N Engl J Med **344:**1140-4.
87. **Wang, L., D. Liu, and P. R. Reeves.** 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:**2598-604.
88. **Wang, L., and P. R. Reeves.** 1994. Involvement of the galactosyl-1-phosphate transferase encoded by the *Salmonella enterica* rfbP gene in O-antigen subunit processing. J Bacteriol **176:**4348-56.
89. **Weadge, J. T., M. M. Palcic, and T. P. Begley.** 2007. Glycosyltransferases, Chemistry of. John Wiley & Sons, Inc.

90. **Weiser, J., Nahm MH.** 2008. Immunity to extracellular bacteria. In: Paul WE ed. Fundamental immunology. 6th edn. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins:1182–1203.
91. **Weiser, J. N., R. Austrian, P. K. Sreenivasan, and H. R. Masure.** 1994. Phase variation in pneumococcal opacity: relationship between colonial morphology and nasopharyngeal colonization. Infect Immun **62:**2582-9.
92. **Weiser, J. N., *et al.*** 2001. Changes in availability of oxygen accentuate differences in capsular polysaccharide expression by phenotypic variants and clinical isolates of *Streptococcus pneumoniae*. Infect Immun **69:**5430-9.
93. **Whitfield, C.** 2006. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. Annu Rev Biochem **75:**39-68.
94. **Whitney, C. G., *et al.*** 2000. Increasing prevalence of multidrug-resistant *Streptococcus pneumoniae* in the United States. N Engl J Med **343:**1917-24.
95. **Winslow, C. E., *et al.*** 1920. The Families and Genera of the Bacteria: Final Report of the Committee of the Society of American Bacteriologists on Characterization and Classification of Bacterial Types. J Bacteriol **5:**191-229.
96. **Xayarath, B., and J. Yother.** 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:**3369-81.
97. **Yother, J.** 2011. Capsules of *Streptococcus pneumoniae* and other bacteria: paradigms for polysaccharide biosynthesis and regulation. Annu Rev Microbiol **65:**563-81.
98. **Yuasa, R., M. Levinthal, and H. Nikaido.** 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:**433-44.
99. **Zhang, Y., and R. M. Miller.** 1992. Enhanced octadecane dispersion and biodegradation by a *Pseudomonas* rhamnolipid surfactant (biosurfactant). Appl Environ Microbiol **58:**3276-82.
100. **Zhang, Z., T. B. Clarke, and J. N. Weiser.** 2009. Cellular effectors mediating Th17-dependent clearance of pneumococcal colonization in mice. J Clin Invest **119:**1899-909.