Coordination of Primer Sugar Synthesis with O-antigen Initiation in Rhizobium Etli CE3

Tiezheng Li
Marquette University

COORDINATION OF PRIMER SUGAR SYNTHESIS WITH O-ANTIGEN INITIATION IN *RHIZOBIUM ETLI* CE3

by

Tiezheng Li, B.S.

A Dissertation Submitted to the Faculty of the Graduate School,
Marquette University,
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy

Milwaukee, Wisconsin

August 2014

ABSTRACT
COORDINATION OF PRIMER SUGAR SYNTHESIS
WITH O-ANTIGEN INITIATION
IN *RHIZOBIUM ETLI* CE3

Tiezheng Li, B.S.

Marquette University, 2014

All organisms synthesize amidosugars, such as *N*-acetylglucosamine (GlcNAc), and deoxysugars, such as fucose. They are found in important polysaccharides and glycoconjugates such as glycoproteins. *N*-acetylquinovosamine (QuiNAc) is both an amido- and a deoxy-sugar. It is found in many examples of an important prokaryotic glycoconjugate, the lipopolysaccharide (LPS) that coats the surface of Gram-negative bacteria. Like N-glycosylation of glycoproteins, LPS has a portion that is synthesized first on a polyprenyl lipid carrier and then transferred to the rest of the molecule.

QuiNAc is believed to initiate the O-antigen portion of LPS of *Rhizobium etli* CE3. Genetic studies identified three genes, *wreV*, *wreU* and *wreQ*, required for the initial steps of O-antigen synthesis in *R. etli* CE3. Based on the predicted roles of the gene products and the theory of polysaccharide biosynthesis, there was a very straightforward prediction of the initial events: WreV catalyzes conversion of UDP-GlcNAc to its 4-keto-6-deoxy derivative, which WreQ reduces to UDP-QuiNAc, followed by transfer of QuiNAc-1-phosphate to the lipid carrier by WreU. However, the LPS structure of *R. etli wreQ* mutants was also consistent with a second, novel possibility.

These two hypotheses were tested by developing assays *in vitro* for each of the predicted enzymes. Two key findings were 1) that WreU catalysis is 30-fold faster with the 4-keto intermediate as substrate than with UDP-QuiNAc, and 2) WreQ catalyzes the reduction the 4-keto sugar to QuiNAc orders of magnitude faster when it is linked to lipid rather than to UDP.

The results strongly support the second hypothesis, outlined as follows: After 4,6-dehydration of UDP-GlcNAc, the 4-keto-6-deoxysugar-1-phosphate moiety is transferred to the lipid carrier, thereby providing the sugar stem for the rest of O-antigen synthesis. Only then is the 4-keto sugar reduced to QuiNAc. The order of enzyme reactions suggests an interesting chemical coordination between the initiation of the O antigen and QuiNAc synthesis. It also includes the unorthodox completion of deoxysugar synthesis on a lipid carrier. Because the presence of QuiNAc and WreQ homologs are highly correlated in bacterial species, this may be the normal mechanism by which QuiNAc becomes part of bacterial polysaccharides

ACKNOWLEDGEMENTS

Tiezheng Li, B.S.

I would like to express my deepest gratitude to my advisor, Dr. Dale Noel, for his excellent guidance, patience and support. I feel extremely privileged to have had the opportunity to develop as a scientist under his mentorship.

I would like to thank my committee members, Dr. Rosemary Stuart, Dr. Pinfen Yang, Dr. Martin St. Maurice and Dr. Krassimira Hristova, for all the comments and inspirations in my research project, also for their careful reading of this dissertation and critical suggestions. My special thanks go to Dr. Evgueni Kovriguine, for his great help with NMR experiments.

Many thanks to people of the Noel lab that I had the honor to work with, Kristylea Ojeda, Zachary Lunak, Laurie Simonds, Sihui Yang and Aaron Mason. I enjoyed every moment that we spent together. I would also like to thank the biology department for creating an excellent atmosphere for doing research and making friends.

I could not have made it this far without the unconditional love and support from my parents. They encourage me to conquer challenges with their own hard work and bravery.

I dedicate this work to my beloved wife, Linlin. I cannot express how much I appreciate her for being with me, always believing in me, sharing my happiness and cheering me up in the bad days. I feel extremely lucky to have her in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3OMe6dTal	3-O-methyl-6-deoxytalose
AADGal	2-acetamido-4-amino-6-deoxygalactose
ABC	ATP-binding cassette
ADHexu	2-acetamido-2,6-dideoxy-D-xylo-4-hexulose
ATP	adenosine triphosphate
Bac	bacillosamine or 2,4-diamino-2,4,6-trideoxy-D-glucose
COSY	correlation spectroscopy
CPSA	capsular polysaccharide A
DGK	diacylglycerol kinase
Fuc	fucose
FucNAc	*N*-acetyl-D-fucosamine or 2-acetamido-2,6-dideoxy-D-galactose
GC	gas chromatography
GC-MS	gas chromatography-mass spectrometry
gCOSY	gradient-enhanced correlation spectroscopy
GalA	galacturonic acid
GalNAc	*N*-acetyl-D-galactosamine or 2-acetamido-2-deoxy-D-galactose
GDP	guanosine diphosphate
GlcNAc	*N*-acetyl-D-glucosamine or 2-acetamido-2-deoxy-D-glucose
Gm	gentamycin
GTase	glycosyltransferase
His_6	six-histidine
Hz	hertz
Kb	kilobase
Kdo	3-deoxy-D-manno-oct-2-ulosonic acid

Km	kanamycin
LLO	lipid-linked oligosaccharide
LPS	lipopolysaccharide
LPS I	lipopolysaccharide composed of lipid A, core, and O-antigen
LPS II	lipopolysaccharide composed of lipid A and core regions
Man	mannose
MeGlcA	glucuronyl methyl ester
MurNAc	*N*-acetylmuramic acid
NDP	nucleotide diphosphate
NMR	nuclear magnetic resonance
NOE	nuclear overhauser effect
NOESY	nuclear overhauser effect spectroscopy
PGN	peptidoglycan
PHPT	polyisoprenyl-phosphate hexose-1-phosphate transferase
PNPT	polyisoprenyl-phosphate *N*-acetylhexosamine-1-phosphate transferase
ppm	parts per million
QuiN	quinovosamine
QuiNAc	2-acetamido-2,6-dideoxy-D-glucose
*RL*3841	*Rhizobium leguminosarum* bv. *viciae* 3841 or (*R. leguminosarum* 3841)
R-LPS	lipopolysaccharide composed of lipid A and core regions
SDR	short chain dehydratase/reductase
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide electrophoresis
S-LPS	lipopolysaccharide composed of lipid A, core, and O-antigen
SR-LPS	lipopolysaccharide that contain lipid A, core, and a truncated O-antigen
sugar-1-P	sugar-1-phosphate

TLC	thin layer chromatography
TM	transmembrane
Tn	transposon
TOCSY	total correlation spectroscopy
UDP	uridine diphosphate
Und-OH	undecaprenol
Und-P	undecaprenyl phosphate
Und-PP	undecaprenyl pyrophosphate

CHAPTER ONE: INTRODUCTION

A. Lipopolysaccharide (LPS)

The surface of Gram-negative bacteria consists of two membrane bilayers sandwiching a thin peptidoglycan (PGN) layer and the periplasm (Fig. 1). The outer leaflet of the outer-membrane bilayer is mainly composed of lipopolysaccharides (LPS). The LPS molecule typically consists of three structural regions (Fig. 1): a hydrophobic lipid A, a non-repeating core oligosaccharide, and a distal polysaccharide called O antigen. Lipid A comprises the outer monolayer of the outer membrane and anchors the sugar moieties. The structure of lipid A is highly conserved among species (1). The core region extends outward from lipid A and provides an attachment site for the O antigen. O antigen is the most surface-exposed portion of LPS. The extent of structural variation in core oligosaccharides is limited within a given species, or even a genus. In contrast, the structure of O antigens varies from strain to strain in the same species (2).

B. The LPS of *Rhizobium etli* CE3

Rhizobium etli CE3, as a member of the α-proteobacteria, has the canonical Gram-negative cell surface. The structures of all three portions of *R. etli* CE3 LPS have been determined (3-5). It is one of the best known LPSs, both structurally and functionally.

Like in many other Gram-negative bacteria, two forms of LPS exist in *R. etli* CE3: the smooth LPS (or LPS I) and the rough LPS (or LPS II). LPS I contains all three structural regions of an LPS while LPS II contains only lipid A and core, lacking the O antigen. The two forms of LPS can be separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by silver staining (6) (Fig. 2). Mutants deficient in O-antigen biosynthesis can be identified by their altered LPS banding pattern. Those lacking O antigen completely have no visible LPS I bands, and the amount of LPS II per cell is increased (Fig. 2).

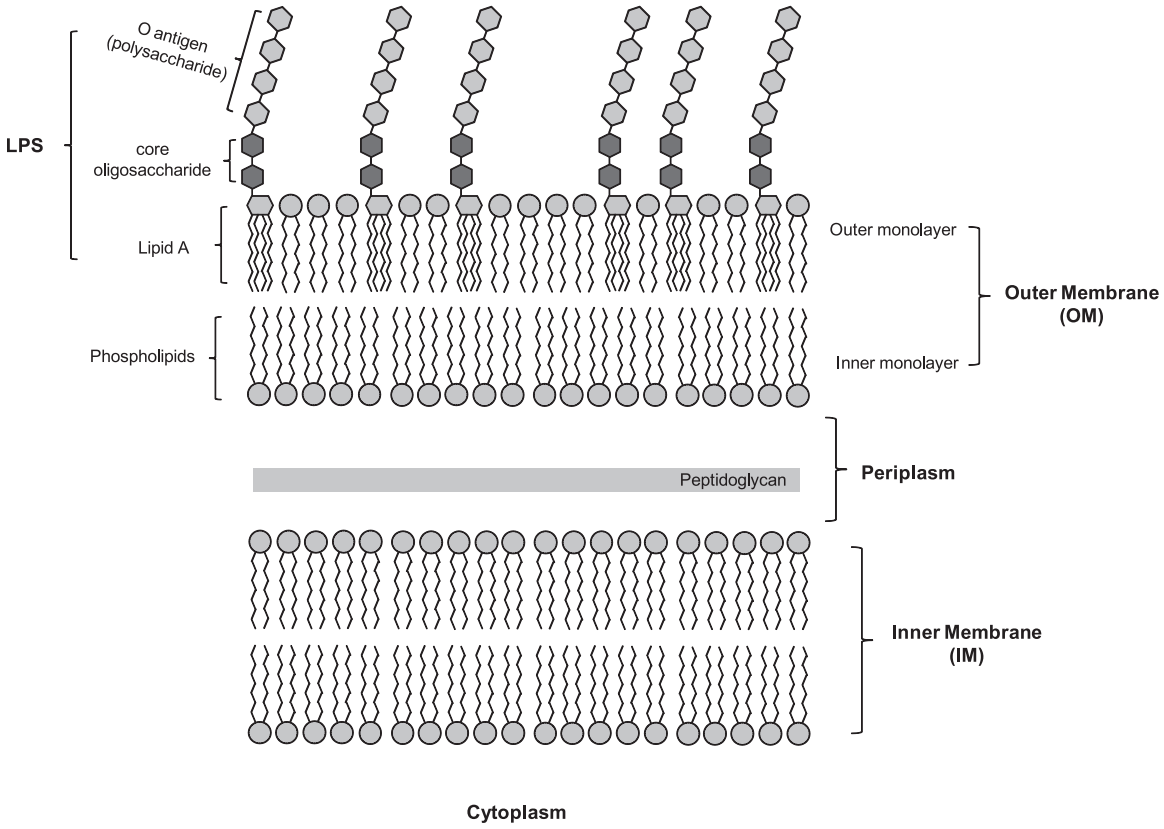


Figure 1. Schematic representation of the surface of gram-negative bacteria. Lipopolysaccharides (LPS) comprise the outer monolayer of the outer membrane. The three domains of LPS, lipid A, core oligosaccharide and O antigen, are shown. The inner monolayer of the outer membrane and the inner membrane bilayer are composed of phospholipids. A thin peptidoglycan layer is located in the periplasm between the two membrane bilayers.

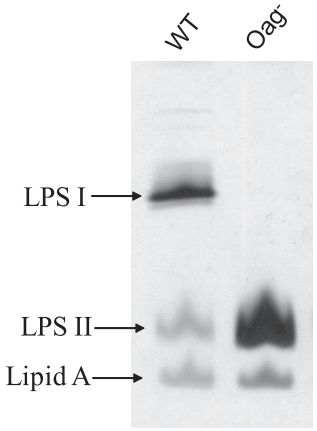


Figure 2. Separation of *R. etli* LPS I and LPS II. A silver-stained SDS-PAGE gel (18%) with LPS samples from whole-cell lysates. *Left lane*, the *R. etli* wildtype strain CE3 makes two forms of LPS, one with O-antigen (LPS I), one without (LPS II); *Right Lane*, an O-antigen mutant (Oag$^-$) strain lacks LPS I, and has more abundant LPS II. This specific Oag$^-$ mutant is defective in the *wreU* gene.

C. The biological role of O antigen and the specific role of *R. etli* CE3 O antigen

The location of O antigen at the cell surface places it at the interface between the bacterium and its surrounding environment. Early interest in the study of O-antigen structure and biosynthesis was stimulated by their roles as essential virulence determinants and their potential application in vaccine development. But now, information is available for a wide range of species with different lifestyles (2). The primary role of the O antigen appears to be protective. Animal pathogens use O antigen to evade host immune responses, particularly the complement system. The assembly of the membrane attack complex is affected by features of O antigen such chain length and relative amount of long-chain LPS (7,8).

In Rhizobium-plant symbiosis, the O antigen may confer resistance to plant defense mechanisms (9) and may serve as signals that trigger the plant to allow the infection to proceed. The O antigens of *Rhizobium* species are essential in the infection of their host legumes to form nitrogen-fixing root nodules (6,10,11). Mutants derived from *R. etli* CE3 that do not make O antigen, make truncated O antigens, or make normal-length O antigen in reduced amount, are deficient in symbiosis (6). The O-antigen mutants were still able to initiate the nodulation process but nodules elicited by them were undeveloped, bacteria-free and unable to fix nitrogen (6).

D. Structure of the *R. etli* CE3 O antigen

The O antigen of *R. etli* CE3 is a fixed-length heteropolymer (4). Its structure has been defined in four parts: the primer, adapter region, repeating units and a terminal residue (Fig. 3) (12). The identity of the first sugar residue of O antigen, as depicted in Fig. 3, is a hypothesis. Only with the biosynthetic analysis reported in this dissertation is this hypothesis now being confirmed. However, strong arguments based on comparison to other O-antigen structures and evidence from previous genetic studies, suggested that the primer sugar is QuiNAc (2-acetamido-2,6-dideoxy-D-glucose or *N*-acetyl-D-quinovosamine) (12). The adapter region refers to sugars

added after the primer and before the repeating units, and it is composed of a mannose (Man) and then a fucose (Fuc) in *R. etli* CE3. The repeating unit of *R. etli* CE3 is a branched trisaccharide composed of a Fuc, a glucuronic acid methyl ester (MeGlcA) and a 3-O-methyl-6-deoxytalose (3OMe6dTal). After this unit is repeated five times the O antigen ends in either a 2,3-di-O-methylated or 2,3,4-tri-O-methylated fucose (4).

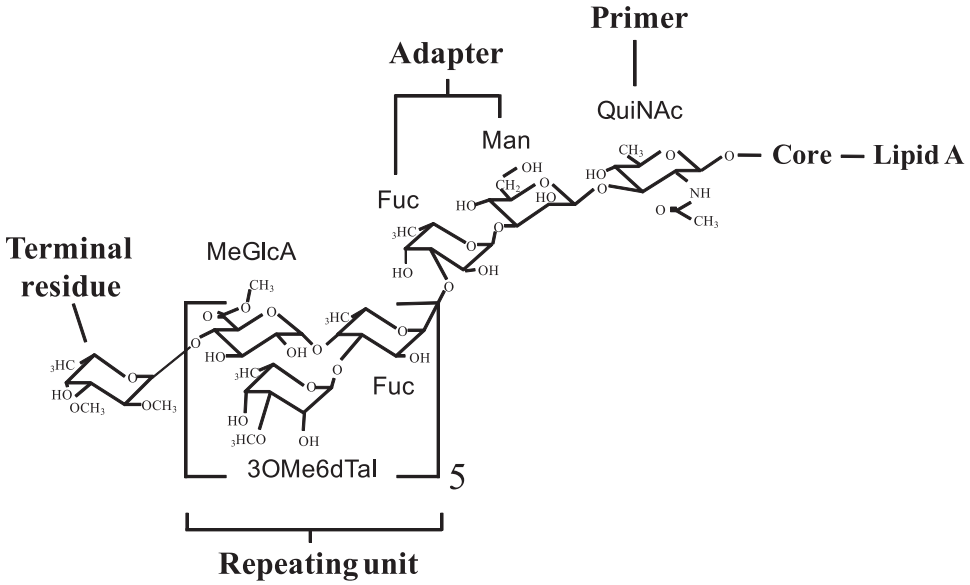


Figure 3. Structure of *Rhizobium etli* CE3 O antigen. The full O-antigen structure is shown linked to the lipid A-core, the structure of which are not shown in detail. The primer, adapter region, repeating units and terminal residue are indicated. The first sugar residue of the O antigen, as depicted here, is a hypothesis. Abbreviations for the sugars: QuiNAc, *N*-acetyl-D-quinovosamine; Man, mannose; Fuc, fucose; MeGlcA, glucuronyl methyl ester; 3OMe6dTal, 3-*O*-methyl-6-deoxytalose; Terminal residue, TOMFuc, 2,3,4-tri-O-methylfucose or DOMFuc, 2,3-di-O-methylfucose.

E. O-antigen biosynthesis

The O-antigen portion of LPS is not synthesized directly on the core oligosaccharide. Instead, it is assembled on a lipid carrier molecule, termed undecaprenyl phosphate (Und-P). This is the same C_{55}-isoprenoid alcohol derivative used for the synthesis of peptidoglycan (13) and most exopolysaccharides (14). During the synthesis phase, the growing polysaccharides are linked at the reducing end to undecaprenyl pyrophosphate (Und-PP) (Fig. 4). Fig. 5 outlines the

overall postulated scheme by which the CE3 O antigen is synthesized on the cytoplasmic face of the inner membrane, exported to the periplasmic face, transferred from Und-PP to lipid A-core, and then translocated to the outer membrane. Despite the accumulated information of hundreds of O-antigen structures and hundreds of biosynthesis gene clusters, there are relatively few detailed biochemical investigations of O-antigen synthesis and the synthetic process remains poorly defined. Many of the individual enzymatic mechanisms are speculative, and even functions based on sequence similarities are tenuous.

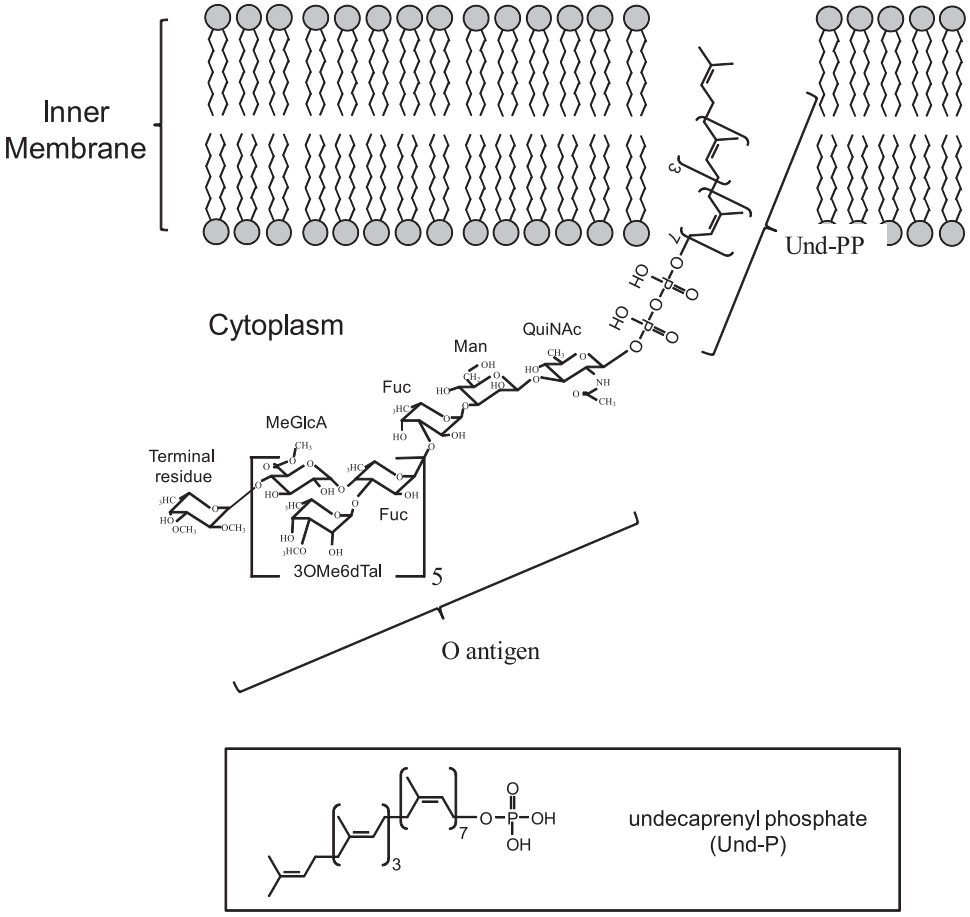


Figure 4. O antigens are assembled on the lipid carrier undecaprenyl phosphate on the cytoplasmic face of the inner membrane. Shown here is the O antigen of *R. etli* CE3 as it putatively would appear at the end of its biosynthesis before transport to the periplasmic face of the inner membrane. *inset*, undecaprenyl phosphate (Und-P).

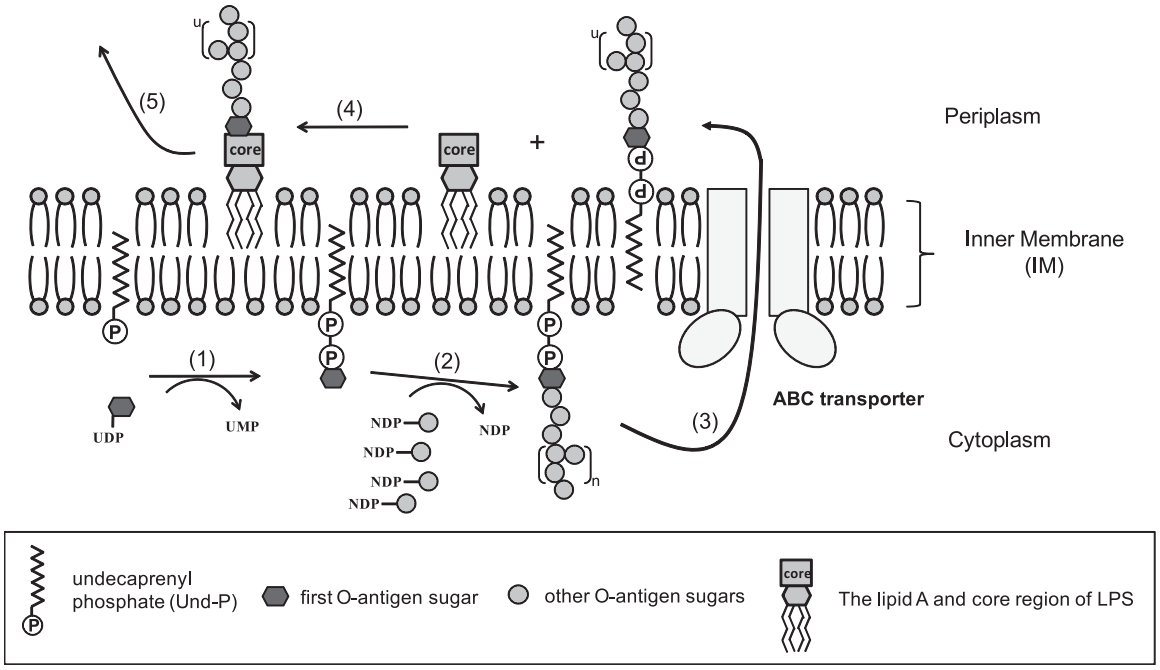


Figure 5. The overall scheme of O-antigen and LPS biosynthesis expected to occur in *R. etli* CE3. The O antigen of *R .etli* CE3 is believed to be synthesized by the ABC-transporter-dependent pathway due to the presence of ABC-transporter genes in its O-antigen cluster (15) and was supported by genetic studies (16). In this pathway, the entire O antigen is assembled on the cytoplasmic face of the inner membrane. The first step is catalyzed by an initiating glycosyltransferase (GTase), which transfers the sugar-1-phosphate moiety from a UDP-sugar donor to the lipid carrier undecaprenyl phosphate (step #1). Chain extension occurs at the nonreducing end by transferring the rest of the O-antigen sugars one after another until completion (step #2). Once completed, the lipid-linked O-antigen intermediate is transported by an ABC transporter to the periplasmic face of the inner membrane (step #3). There, O antigen is ligated to the lipid A and core regions (step #4). Finally, the completed LPS is mobilized to the outer membrane (step #5). The focus of this dissertation is on the initial steps. The purpose of this figure is to put these steps into the context of the overall process. In this figure, no attempt is made to depict the biosynthetic enzymes or their possible interactions with each other and the ABC transporter. The other major pathway, Wzy-dependent, differs from this pathway in that each repeating unit is synthesized on a lipid carrier on the cytoplasmic face and transported to the periplamic face of the inner membrane individually. There they are polymerized in a way that chain extension is at the reducing end, like peptidoglycan.

F. Genetic studies of *R. etli* CE3 O-antigen biosynthesis

In general, O-antigen synthesis genes are found in clusters (2,17). Those in *R. etli* CE3 are no exception. Nearly all genes required to synthesize the known *R. etli* CE3 O-antigen structure are located on the chromosome, spanning nucleotides 784,527 to 812,262 (Fig. 6A) (15) (previously designated the *lps* α region (18)). This gene cluster includes genes encoding eight

putative glycosyltransferases (GTases), the ABC-transporter proteins, enzymes for sugar synthesis or modification and several genes with unknown function (12). In addition to this cluster, genes required for O-antigen biosynthesis have been identified elsewhere on the chromosome (Fig. 6B) (19) and on plasmid pCFN42b (Fig. 6C) (20). Except for well-known orthologous genes, genes involved in O-antigen biosynthesis in *R. etli* have been given the designation *wre* ("*w*" stands for O-antigen genes, "*re*" stands for *R. etli*) (21), in order to conform with the conventions of naming O-antigen-related genes (22).

The functions of most of the genes have been assigned based mainly on mutagenesis studies and sequence homology of the encoded proteins to well-characterized enzymes. For example, each putative GTase gene has been studied individually by mutagenesis and their predicted functions can account for each step of O-antigen biosynthesis (12). The availability of complete O-antigen structure and genetic information make *R. etli* CE3 one of the best candidates for detailed biochemical investigations of heteropolymeric O-antigen biosynthesis. Gene *wreU*, *wreV* and *wreQ* will be the focus of this dissertation and proved to be responsible for the synthesis and attachment of the first O-antigen sugar, QuiNAc. Interestingly, all of them are located outside the 28-kb O-antigen genetic cluster.

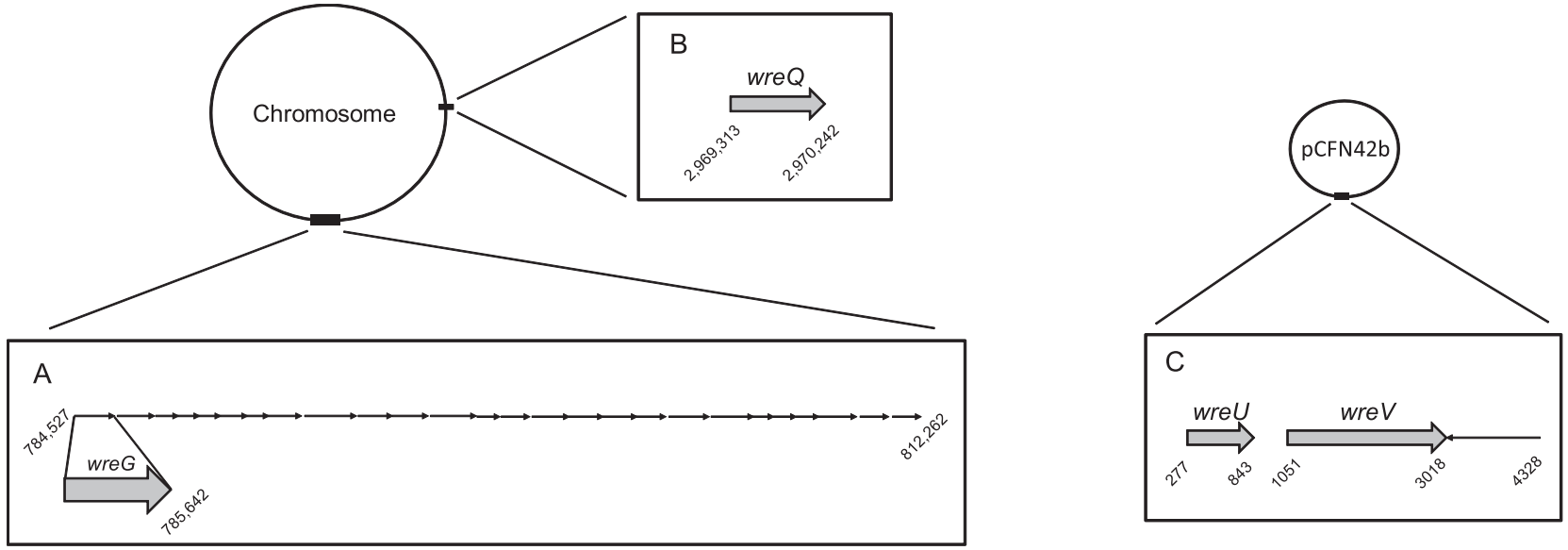


Figure 6. *R. etli* CE3 O-antigen genetic clusters. (A) A 28-kilobase cluster (previously called *lps* region α) on the chromosome, spanning nucleotides 784,527 to 812,262 of the genome sequence and consisting of 25 predicted ORFs. (B) A single ORF spanning nucleotides 2,969,313 to 2,970,242 on the chromosome, that has been previously identified as being required for O-antigen synthesis (19). (C) A 4-kilobase cluster on plasmid pCFN42b consisting of three predicted ORFs. The genes specifically studied in this dissertation, *wreG*, *wreQ*, *wreU* and *wreV*, are shown with enlarged arrows, with name on top. Nucleotide numbering indicates the start and stop of each of the four genes and the edge of each genetic cluster.

G. The initial step of O-antigen biosynthesis

To the extent studied previously, all O-antigen biosyntheses have similar initiation reactions (2). An initiating GTase catalyzes the formation of an Und-PP-linked glycose by transfer of a sugar-1-phosphate (sugar-1-P) residue to Und-P. This enzyme differs from other O-antigen GTases in two aspects (Fig. 7): first, it recognizes a hydrophobic lipid carrier rather than the sugar acceptor used by the other GTases; second, it transfers the sugar-1-P of a nucleotide diphosphate (NDP) sugar donor whereas other GTases only transfer the sugar.

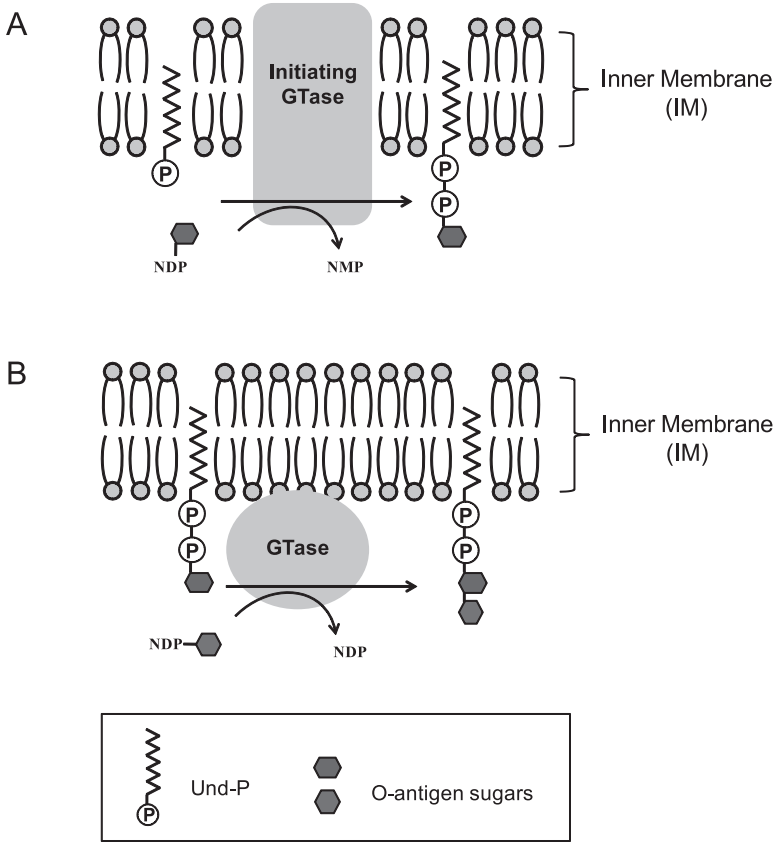


Figure 7. Comparing the reactions catalyzed by the initiating GTase and other GTases in O-antigen biosynthesis. (A) All characterized initiating GTases are integral membrane proteins. They catalyze the transfer of sugar-1-phosphate from a UDP-sugar donor to the lipid carrier Und-P. (B) Most other GTases involved in O-antigen synthesis are not predicted to have transmembrane segments. They transfer only the sugar moiety from an NDP-sugar donor to another sugar acceptor.

Biochemical characterizations of the initiating GTases are relatively limited, mainly due to problems encountered with their overexpression and purification. The best characterized ones are WecA from *E. coli* (formerly Rfe) (23) and WbaP from *Salmonella enterica* (formerly RfbP) (24).

E. coli WecA is an integral membrane protein with 11 transmembrane (TM) segments (Fig. 8A) (25). Recently, the WecA protein from the thermophilic bacterium *Thermotoga maritima* was successfully purified and shown to catalyze the transfer of GlcNAc-1-P moiety from UDP-GlcNAc onto the lipid carrier Und-P (26).

WbaP initiates O-antigen synthesis in *Salmonella enterica* by catalyzing the transfer of galactose-1-P onto Und-P. The WbaP protein possesses three predicted domains: an N-terminal region containing four predicted transmembrane segments, a large central periplasmic loop and a C-terminal domain containing the last transmembrane segment and a large cytoplasmic tale (Fig. 8B) (27,28). The C-terminal domain alone is sufficient for the GTase activity of WbaP; the truncated protein ($WbaP_{250\text{-}476}$) could complement the synthesis of O antigen in a *S. enterica* Δ*wbaP*::*cat* mutant strain (27). Recently, a truncated WbaP containing the C-terminal domain ($WbaP_{258\text{-}476}$) was purified with a removable N-terminal thioredoxin fusion and shown to possess GTase activity *in vitro* (29).

H. The initial step of O-antigen biosynthesis in *R. etli* CE3

The initiating GTase—There is strong, albeit indirect, evidence that the gene *wreU* encodes the initiating GTase for *R. etli* CE3 O-antigen biosynthesis. The argument is based on sequence homology, *wreU* gene mutagenesis, and chemical analysis of the LPS produced by the *wreU* mutant (12). The WreU protein is predicted to possess a single TM segment at its N-terminus, followed by a large cytoplasmic tail (Fig. 8C). Sequence alignment showed that WreU is homologous to the C-terminal domain of WbaP (Fig. 9).

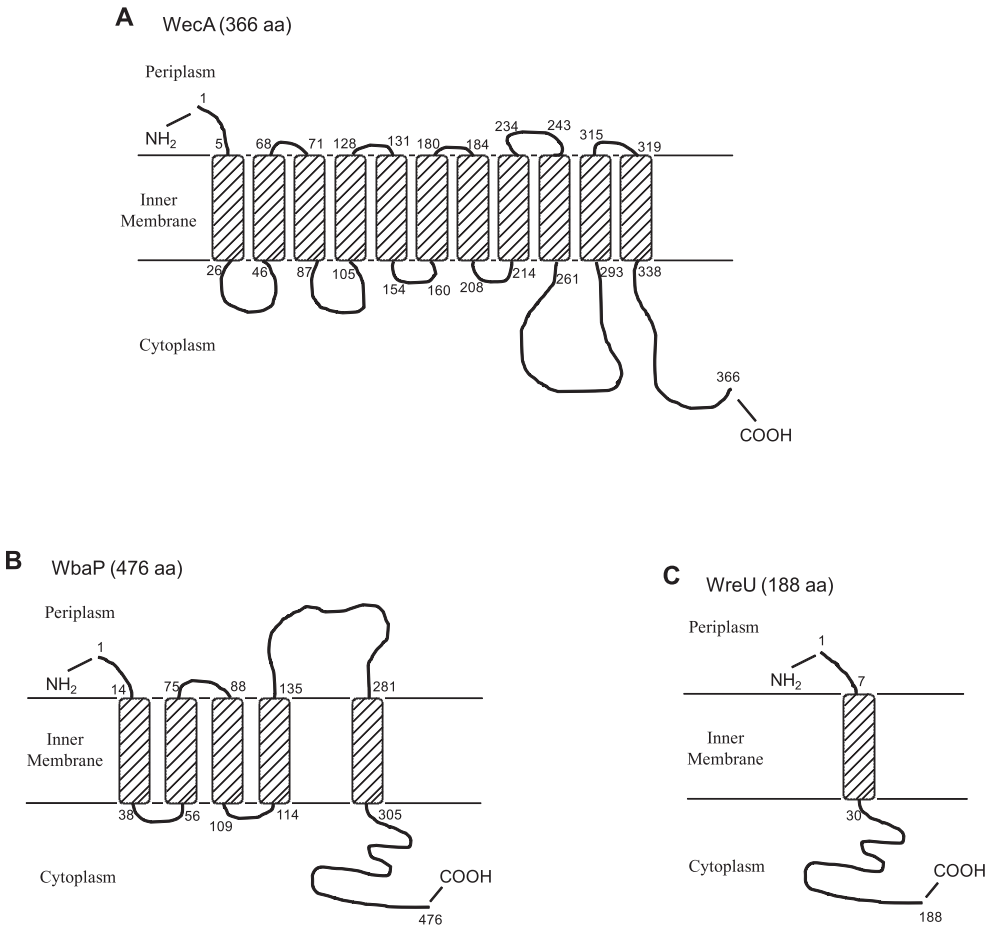


Figure 8. Topological models of initiating GTases. The shaded rectangular boxes represent transmembrane (TM) segments. The numbers indicate the amino acid positions of the boundary of each TM segment. (A) *Escherichia coli* WecA; (B) *Salmonella enterica* WbaP; (C) *Rhizobium etli* WreU.

The first sugar of O antigen—As mentioned above, the first sugar (primer) of the *R. etli* CE3 O antigen was predicted to be QuiNAc. This prediction was based on comparison to other O-antigen structures and was supported by genetics studies (12). LPS sugar composition analysis of many O-antigen mutants showed that they each had significantly less QuiNAc than galactose, a core-specific sugar, suggesting that QuiNAc is not synthesized as part of the core region (12). Studies of *R. etli* CE3 LPS showed that QuiNAc is linked to the non-reducing end of a Kdo (3-deoxy-D-manno-oct-2-ulosonic acid) residue (4,31). Biosynthetic/genetic studies strongly supported that this Kdo is part of the biosynthetic core of the LPS (5,32-34).

```
WreU      ---MGLKRAVDFLLALIASVILLVPILVVALSVRLTSPGPILYWSKRIGRFNQIFLMPKF   57
WbaP      RSSRFLKRTFDIVCS-IMILIIASPLMIYLWYKVTRDGGPAIYGHQRVGRHGKLFPCYKF   332
               ***:.*:: : *  :*:  *:::        . ** :*  :*:**. ::*   **

WreU      RSMRVDTPTVATHLLEN---------------PERFLTPIGSFLRKSSLDELPQLWCILA   102
WbaP      RSMVMNSQEVLKELLANDPIARAEWEKDFKLKNDPRITAVGRFIRKTSLDELPQLFNVLK   392
          *** :::  * ..** *                :  :* :* *:**:********: :*

WreU      GKMSFVGPRPALYNQYDLIELRTVYGVDKLLPGLTGWAQINGRDELPIPEKVKFDVEYLE   162
WbaP      GDMSLVGPRPIVSDELERYCDDVDY-YLMAKPGMTGLWQVSGRNDVDYDTRVYFDSWYVK   451
          *.**:***** : :: :     . *      **:**  *:.**:::    :* **  *::

WreU      RRSLGFDMRILFLTAEKVVRRKGIKH    188
WbaP      NWTLWNDIAILFKTAKVVLRRDGAY-    476
          . :*  *: *** **: *:**.*
```

Figure 9. Alignment of the predicted WreU amino acid sequence with the C-terminal domain of WbaP (amino acids 274-476) by use of the Clustal Omega program (30).

The initiating reaction—Based on these independent evidence, the simplest hypothesis is that WreU transfers the QuiNAc-1-P to the lipid carrier Und-P to form undecaprenyl-pyrophosphate-linked QuiNAc (Und-PP-QuiNAc), the first intermediate in O-antigen biosynthesis.

However, these inferences above remain to be verified experimentally *in vitro*. In particular, no enzyme has been directly demonstrated to be an initiating GTase that transfers QuiNAc-1-P. One obstacle to such studies is that the expected biosynthetic donor of the QuiNAc residue, UDP-QuiNAc, has never been chemically synthesized *in vitro*.

I. The proposed two-step biosynthetic pathway for UDP-QuiNAc

The proposed biosynthetic pathway of UDP-QuiNAc in *R. etli* CE3 involves two steps (19,35) (Fig. 10). In the first step, the precursor UDP-GlcNAc is converted by a 4,6-dehydratase to its 4-keto-6-deoxy derivative, UDP-2-acetamido-2,6-dideoxy-D-xylo-4-hexulose (UDP-ADHexu). In the second step, UDP-ADHexu is further converted to UDP-QuiNAc by a 4-reductase.

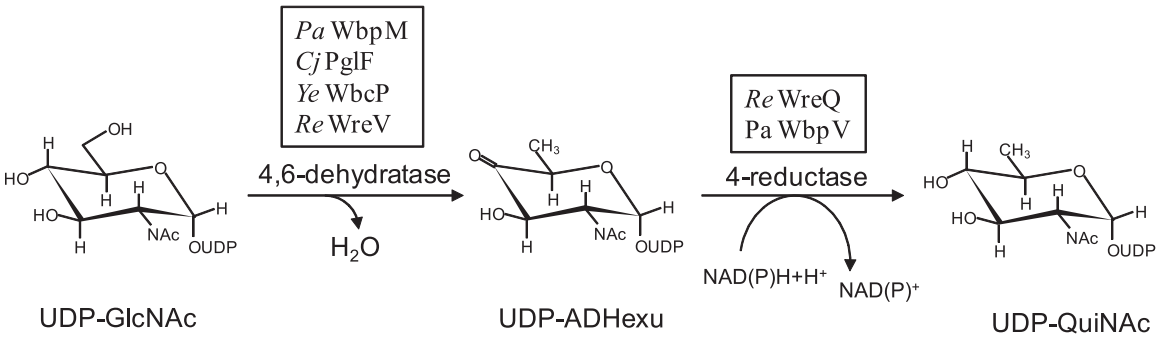


Figure 10. The suggested biosynthetic pathway for the synthesis of UDP-QuiNAc. Enzymes that has been demonstrated biochemically (such as *Pa* WbpM, *Cj* PglF and *Ye* WbcP) or has been proposed (such as *Re* WreV, *Re* WreQ and *Pa* WbpV) to catalyze each reaction are listed in rectangle boxes. *Pa*, *Pseudomonas aeruginosa*; *Cj*, *Campylobacter jejuni*; *Ye*, *Yersinia enterocolitica*; *Re*, *Rhizobium etli*.

The first step is shared in the biosynthetic pathways of many UDP-6-deoxy-sugars derived from UDP-GlcNAc (36) and this step has been reproduced *in vitro* by use of several homologous enzymes, including WbpM from *Pseudomonas aeruginosa* (37), PglF (Cj1120c) from *Campylobacter jejuni* (38,39) and WbcP from *Yersinia enterocolitica* (40). All these proteins are integral transmembrane proteins and they belong to a group of long-chain dehydratases characterized by the NAD-binding domains (G*XX*G*XX*G) and the SMK catalytic triad (41). The *wreV* gene of *R. etli* CE3 encodes a protein that is highly homologous to WbpM (Fig. 11A). Thus, WreV may catalyze the conversion of UDP-GlcNAc to UDP-ADHexu in *R. etli* CE3.

In contrast, the second-step reaction and an enzyme responsible it have never been demonstrated *in vitro*. The *wreQ* gene of *R. etli* CE3 has been proposed to encode the 4-reductase for the second-step of UDP-QuiNAc synthesis (19). The predicted amino acid sequence suggests that WreQ belongs to the short chain dehydratase/reductase (SDR) superfamily, characterized by the G*XX*G*XX*G nucleotide-binding motif and the SYK catalytic triad (19,42).

WreQ is homologous to the predicted protein sequence encoded by the *wbpV* gene of *P. aeruginosa* serotype O6 (Fig. 11B). *wbpV* gene and the aforementioned *wbpM* gene encoding the 4,6-dehydratase are both located in a cluster of genes involved in B-band O-antigen synthesis in *P.*

aeruginosa O6. They have have been independently proposed to catalyze UDP-QuiNAc synthesis (43).

A

```
WreV      ----------MPRAAKRALALLVDSGLCVLTIWLAYCFRLNEWT---VLTGVQWLPVFV      46
WbpM      MLDNLRIKLLGLPRRYKRMLQVAADVTLVWLSLWLAFLVRLGTEDMISPFSGHAWLFIAA      60
                     :**  ** * : .*  *  *::***: .**        ::*  ** : .

WreV      SLCMALPIFIVMGMYRAIFRYANMAAFIAVLKAIAIYGVAFMTIFT--ALSVPGVPRTVG      104
WbpM      -PLVAIPLFIRFGMYRAVMRYLGNDALIAIAKAVTISALVLSLLVYWYRSPPAVVPRSLV      119
             :*:*:** :*****::**    *:**: **::* .:.:  :.         ***::

WreV      ILQPFLLLIAIGLSRLGI-RYWLGDTYQRI------LHQNTLAKVLIYGAGNAGRQLAGA      157
WbpM      FNYWWLSMLLIGGLRLAMRQYFMGDWYSAVQSVPFLNRQDGLPRVAIYGAGAAGNQLVAA      179
          :   :* :: **  **.: :*::** *. :       :*: * :* ***** **.**..*

WreV      LTNSAELNVVGYLDDDPRLKGGIMGGLPIYDPSDLPVLAETLGVHNVLLALPSASRQRRN      217
WbpM      LRLGRAMRPVAFIDDDKQIANRVIAGLRVYTAKHIRQMIDETGAQEVLLAIPSATRARRR      239
          *  .  :. *.::*** ::   ::.** :*  ..:  : :  *.::****:***:* **.

WreV      EILEHIRKARVNVRTLPDLTALAQGRVAVSDIRELEIEDLLGREAVAPRQELLDKAMRNK      277
WbpM      EILESLEPFPLHVRSMPGFMDLASGRVKVDDLQEVDIADLLGRDSVAPRKELLERCIRGQ      299
          **** :.   :.**::* :  **.*** *.*::*::* *****::****:***::.:* :

WreV      VVMVTGAGGSIGGELCRQILRNAPSSLILIDQNEFALYNIHAELLKLAEVYKHEGLQIVP      337
WbpM      VVMVTGAGGSIGSELCRQIMSCSPSVLILFEHSEYNLYSIHQELERRIK-RESLSVNLLP      358
          ************.******:  :** ***:::.*: **.** ** :  :  :  .::::*

WreV      ILCSVRDQDRMEHIMESWRPQTLYHAAAYKHVPLVEHNAVEGIKNNVMGTLITARAANKY      397
WbpM      ILGSVRNPERLVDVMRTWKVNTVYHAAAYKHVPIVEHNIAEGVLNNVIGTLHAVQAAVQV      418
          ** ***: :*: .:*.:*: :*:**********:**** .**: ***:*** :.:** :

WreV      GVSNFVLISTDKAVRPTNVMGASKRLAEMVLQALAAELTA----------DRLRTNFSMV      447
WbpM      GVQNFVLISTDKAVRPTNVMGSTKRLAEMVLQALSNESAPVLFGDRKDVHHVNKTRFTMV      478
          **.******************::**********: * :           .  :*.*:**

WreV      RFGNVLGSSGSVVPLFRQQIKEGGPVTLTHPKITRYFMTISEASQLVIQAGAMAEGGDVF      507
WbpM      RFGNVLGSSGSVIPLFREQIKRGGPVTVTHPSITRYFMTIPEAAQLVIQAGSMGQGGDVF      538
          ************:****:***.*****:***.******** **:*******:*.:*****

WreV      LLDMGEPVRIADLARKMVELSGLTVRDVDNPEGDIELAVTGLRPGEKLYEELLIGDNPET      567
WbpM      VLDMGPPVKILELAEKMIHLSGLSVRSERSPHGDIAIEFSGLRPGEKLYEELLIGDNVNP      598
          :**** **:* :**.**:.****:**.  .*.*** : .:*****************  :

WreV      TEHPRIMKAREDFLFWAELSKKLNSLNASLDRNDMVAARATLAELVSGYSSTGEVSDLAF      627
WbpM      TDHPMIMRANEEHLSWEAFKVVLEQLLAAVEKDDYSRVRQLLRETVSGYAPDGEIVDWIY      658
          *:** **:*.*:.* *  :.  *:.* *:::::*   .*  * * ****:  **: *  :

WreV      TGAETNPAHEDVNALASGAGNQLRTTGA     655
WbpM      RQRRREP--------------------      665
             . :*
```

Figure 11. (A)

B

```
WreQ      ---MRCLVTGAAGFVGSPLVKRLHAEKIYDLVATTRSQTPAFPPEVAHFPIEITGGTDWT      57
WbpV      MTRHNVLVTGATGFIGAALVNSLCSSGQYKVWAGCRRRGGAWPRGVTPLLLGELGSSVVW      60
              . *****:**:*: **: * :.  *.: *  * :  *:*  *: : :   *.:

WreQ      AALEGVDVIVHLAARVHIMNDRAADPLAEFRRTNTAAALNLAEQAASAGVKRFVFVSTIK      117
WbpV      DAESAIDTVVHCAARVHVMSETASDPLVEFRKANVQGTLDLAREAVSRGVRRFIFISSIK      120
           * ..:*.:** *****:*.: *:***.***::*. .:*:**.:*.* **:**:*:*:**

WreQ      VNGE--ENDRPFRHDDRPKPIDPYGISKLECEIGLREIAARTGMEVVIIRPPLVYGPGAR      175
WbpV      VNGEGTEPGRPYTADSPPNPVDPYGVSKREAEQALLDLAEETGLEVVIIRPVLVYGPGVK      180
          ****  *  **:  *. *:*:****:** *.* .* ::* .**:******* ******.:

WreQ      GNFALLVNLVRKKLPLPFASLKNHRTLVAVQNLVDLIIACATHPAAPGEIFLAGDGEDLS      235
WbpV      ANVQTMMRWLKRGVPLPLGAIHNRRSLVSLDNLVDLIITCIEHPAAVGQVFLVSDGEDLS      240
          .*.  ::. ::: :***:.:::*:*:**:::*******:*  **** *::**..******

WreQ      TPALIRGIAAGLGVKPMLVPFPPALLQMAAKALGKEAVYQRLCGSLQVDITRARDVLGWS      295
WbpV      TTELLRRMGRALGAPARLLPVPASWIGAAAKVLNRQAFARRLCGSLQVDIMKTRQVLGWT      300
          *  *:* :. .**.   *:*.* : :  ***.* ::*. :********** ::*:****:

WreQ      PVVTPREGLKLAVE------    309
WbpV      PPVGVDQALEKTARSFLDRQ    320
          * *   :.*: :..
```

Figure 11. Clustal alignments of *R. etli* protein WreV and WreQ with homologs in *Pseudomonas aeruginosa*. Shaded areas indicate characteristic motifs. (A) Alignment of the predicted protein sequence of WreV (655 aa) with WbpM (665 aa) of *Pseudomonas aeruginosa* PAO1 (O5). WbpM is a characterized UDP-GlcNAc 4,6 dehydratase It belongs to a group of long-chain dehydratases that catalyze the same reaction and are characterized by the G*XX*G*XX*G motif for NAD(P) binding and the group-specific SMK catalytic triad. WreV belongs to this group by sequence similarity. (B) Alignment of the predicted protein sequence of WreQ (309 aa) and WbpV (320 aa) of *P. aeruginosa* O6. Both proteins belong to the short chain dehydratase/reductase (SDR) superfamily. They possess the G*XX*G*XX*G motif and the typical SYK catalytic triad found in most SDR enzymes.

J. *R. etli wreQ* mutant and suppression of this mutant

A *R. etli* mutant strain (CE166) carrying a Tn*5* transposon in the *wreQ* gene locus exhibits an altered LPS phenotype when analyzed by SDS-PAGE. CE166 (*wreQ*::Tn*5*) makes LPS I (O-antigen containing LPS) of normal size, but the abundance is only about 40% of that in the wildtype strain (CE3) (35). Characterization of the O antigen from CE166 showed that QuiNAc was replaced by its 4-keto derivative ADHexu, which is the proposed intermediate in QuiNAc synthesis (19). This observation led to the inference that *wreQ* encodes the 4-reductase for reducing ADHexu to QuiNAc (19).

Further chemical analysis of sugar linkage revealed that the ADHexu residue in CE166 was located at the same position as QuiNAc in the CE3 LPS. Based on the prediction that QuiNAc is the first sugar of the O antigen in CE3, it is reasonable to assume that ADHexu is the first O antigen sugar in the CE166. The fact that CE166 makes an O antigen of normal length indicates that, first, the initiating GTase (putatively WreU) must be able to transfer the ADHexu-1-P intermediate, as the authentic QuiNAc-1-P, to the lipid carrier Und-P; second, the GTase that catalyzes the second step of O-antigen biosynthesis must be able to recognize ADHexu as an acceptor for attaching the second sugar, in addition to the proposed acceptor QuiNAc.

However, the reduced amount of LPS I in CE166 suggests that ADHexu is less efficient in the catalysis. It is possible that the activity of one or more enzymes involved in O-antigen biosynthesis is affected by the difference between ADHexu and QuiNAc. Candidates include the initiating GTase (putatively WreU) and the GTase for the subsequent catalytic steps.

The amount of LPS I can be restored to nearly wildtype level by introducing into strain CE166 recombinant plasmids that contain a 7.8-kb stretch of the *R. etli* CE3 *lps* genetic region α even though this suppressing DNA does not carry the *wreQ* gene (35). However, the QuiNAc residue is still replaced in the resulting strain by ADHexu (19,35). So, the phenomenon was referred to as extragenic suppression, but not complementation. The mechanism of the suppression is unclear. One possibility is that the suppression is due to the multicopy dosage of a gene or genes in the 7.8-kb DNA fragment that promote the catalysis.

The *R. etli wreQ*::Tn*5* mutant strain CE166 provides a useful tool for *in vivo* study of the initial step of O-antigen biosynthesis, especially for identifying the gene products affected by the absence of QuiNAc.

```
WreG      MRVLHFFKTYWPDTFGGVERTIHAIAKGTQKHGICSEVLSLSPSPDTNTMLFDGHMAYKA      60
WbdC      MRVLHVYKTYYPDTYGGIEQVIYQLSQGCARRGIAADVFTFSPDKDTGPVAYEDHRVIYN      60
          *****.:***:***:**:*:.*: :::*  ::**.::*::::**. **  : :: * .

WreG      KLDFEFASTGFSRDVFRHFRKLSKEADIIHFHFPWPLMDIVDLATRHGKPAIVTYHSDIV      120
WbdC      KQLFEIASTPFSLKALKRFKLIKDDYDIINYHFPFPFMDMLHLSARPDARTVVTYHSDIV      120
          *  **:*** ** ..::::*: :..: ***.:***:*:**::.*::*    ::********

WreG      KQKTLLKFYRPLMSRFLRGMDRIVATSPNYLQSSEVLRAFSSKTSVIPLGLSEGDYPVVA      180
WbdC      KQKRLMKLYQPLQERFLSGVDCIVASSPNYVASSQTLKKYLDKTVVIPFGLEQQDVQHD-      179
          *** *:*:*:** .*** *:* ***:****: **:.*: : .** ***:**.: *

WreG      EEAKARWRSRLPKRFFLFIGVLRYYKGVHVLLEAARQTGFDVVLIGNGPTEDELKRTVRQ      240
WbdC      PQRVAHWRETVGDKFFLFVGTFRYYKGLHILMDAAERSRLPVVVVGGGPLESEVRREAQQ      239
           :  *:**. : .:****:*.:*****:*:*::**.:: : **::* ** *.*::* .:*

WreG      QNLSNVHFLGGLPDADKVAILDLSLGLVFPSHLRSEAFGLSLVEAAMRGKPMISCEIGTG      300
WbdC      RGLSNVMFTGMLNDEDKYILFQLCRGVVFPSHLRSEAFGITLLEGARFARPLISCEIGTG      299
          : **** * * * * **  :::*. *:************::*:*.*  .:*:********

WreG      TSFVNLDGETGLVIPPNDTTALSQAMARLWNDDPLAISYGVNARQRYEVNFTAERMAQSY      360
WbdC      TSFINQDKVSGCVIPPNDSQALVEAMNELWNNEETSNRYGENSRRRFEEMFTADHMIDAY      359
          ***:* *  :* ******: ** :** .***::  :  ** *:*:*:*  ***::* ::*

WreG      AELYRNVAGSR-    371
WbdC      VNLYTTLLESKS    371
          .:** .:  *:
```

Figure 12. Clustal alignment the predicted sequence of *R. etli* CE3 protein WreG with WbdC of *E. coli* serotype O8. WbdC has been characterized *in vitro* to be a mannosyltransferase. It catalyzes the transfer of mannose from GDP-mannose to the GlcNAc residue of a GlcNAc-PP-lipid acceptor.

K. Addition of the second sugar residue in O-antigen biosynthesis in *R. etli* CE3

In the O-antigen of *R. etli* CE3, a mannose is attached to the non-reducing end of QuiNAc (12). The gene encoding the GTase for adding this mannose residue has been assigned to *wreG*, which is located in the chromosomal O-antigen genetic cluster (12). WreG is homologous to the mannosyltransferase WbdC of *E. coli* O8 and O9a (Fig. 12), which has been shown to add mannose to the GlcNAc moiety of a lipid-linked GlcNAc acceptor *in vitro* (44). In *R. etli* CE3, WreG may catalyze the second step of O-antigen biosynthesis by transferring mannose to the QuiNAc moiety of Und-PP-QuiNAc, the predicted first O-antigen biosynthesis intermediate.

The success of this step should depend on whether the preceding steps have occurred normally. Therefore, it can be used to validate the *in vitro* investigation of the initial step(s).

L. Hypotheses and Aims

The overall goal of this dissertation is to elucidate the initial step of O-antigen biosynthesis in *R. etli* CE3. This initial step includes not only the initiation reaction catalyzed by the initiating GTase, but also reactions that lead to the synthesis of the putative first sugar QuiNAc. Evidence from previous genetic studies and from sequence similarities to other proteins could not provide a definitive answer for how O-antigen biosynthesis is initiated in *R. etli* CE3. Specifically, the exact reactions and the order in which the involved enzymes exert their activity remain to be defined. Based on the available information, two hypotheses are proposed and will be tested by a combination of *in vitro* and *in vivo* approaches (Fig. 13 and Fig. 14).

Hypothesis 1: The nucleotide sugar donor of QuiNAc is UDP-QuiNAc, which is synthesized from UDP-GlcNAc following the proposed two-step pathway (Fig. 10). WreQ catalyzes the second step by reducing UDP-ADHexu to UDP-QuiNAc. WreU initiates O-antigen biosynthesis by transferring QuiNAc-1-P from UDP-QuiNAc to Und-P, resulting in the first O-antigen intermediate Und-PP-QuiNAc, which then serves as the acceptor for the second sugar mannose, transferred by WreG (Fig. 13A and Fig. 14A).

Hypothesis 2: After UDP-ADHexu is synthesized following the first step of the proposed synthesis pathway for UDP-QuiNAc (Fig. 10), WreU initiates O-antigen biosynthesis by transferring ADHexu-1-P from UDP-ADHexu to Und-P. The resulting Und-PP-ADHexu is reduced by WreQ to Und-PP-QuiNAc, which is the same result of the initial steps as in hypothesis 1. From that point onward, the remaining steps of O-antigen biosynthesis would be the same in both hypotheses (Fig. 13B and Fig. 14B).

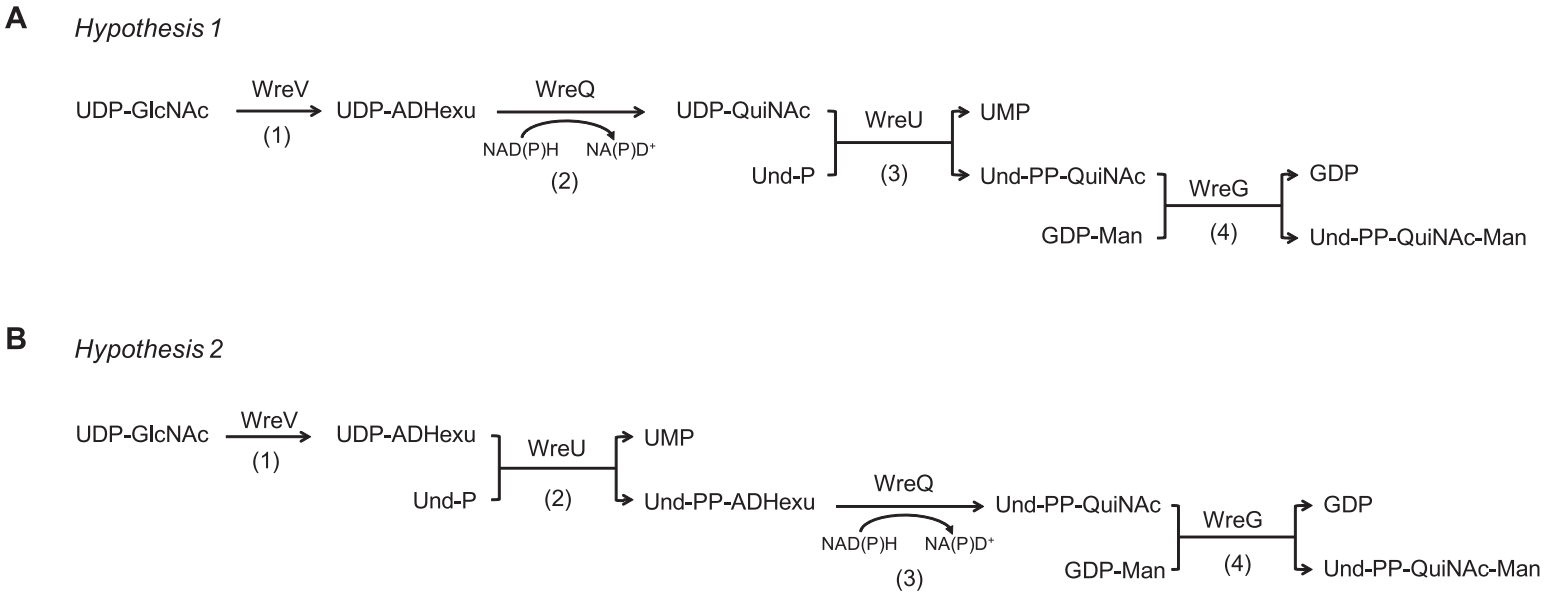


Figure 13. The two hypotheses of O-antigen initiation in *R. etli* CE3. Reactions in each hypothesis and the enzyme that catalyzes each reaction are indicated. In both hypotheses, reactions (1) and (4) are the same. Reaction (1) is the conversion of UDP-GlcNAc to UDP-ADHexu, catalyzed by the predicted 4,6-dehydratase WreV. Reaction (4) is the addition of the second O-antigen sugar mannose to the first O-antigen intermediate Und-PP-QuiNAc. The two hypotheses differ in reactions (2) and (3). (A) In hypothesis 1, ADHexu is reduced to QuiNAc on the UDP linkage by WreQ and then QuiNAc-1-P is transferred by WreU. (B) In hypothesis 2, ADHexu-1-P is transferred by WreU first and then ADHexu is reduced to QuiNAc by WreQ on the Und-PP linkage.

Hypothesis 1 was favored initially because, first, it includes the proposed biosynthetic pathway of UDP-QuiNAc; second, it allows QuiNAc synthesis to occur on a nucleoside diphosphate linkage in all steps, which is the convention for most sugar synthesis (36). It is highly unusual, if not unprecedented, for sugar synthesis to be completed after transfer of an intermediate form to a lipid, as proposed in hypothesis 2. However, both hypotheses can account for the phenotype of strain CE166 (*wreQ*::Tn*5* mutant).

Under hypothesis 1, CE166 phenotype may be explained as follows. Although the preferred substrate of WreU is UDP-QuiNAc, it has a relaxed specificity and can use UDP-ADHexu as substrate in CE166 in which UDP-ADHexu cannot be reduced to UDP-QuiNAc. However, the activity of WreU is much lower with UDP-ADHexu, which leads to slower O-antigen initiation reaction and eventually a lower amount of LPS I.

Under hypothesis 2, CE166 phenotype is explained differently, as follows. Because WreU uses UDP-ADHexu as the substrate to initiate O-antigen biosynthesis, its activity should not be affected in the *wreQ* mutant. However, WreG may greatly prefer the acceptor Und-PP-QuiNAc rather than Und-PP-ADHexu for attaching the second O-antigen sugar mannose. Because only Und-PP-ADHexu is available as an acceptor substrate in CE166 (not being able to convert it to Und-PP-QuiNAc), WreG may catalyze mannose addition so slowly that it limits the rate of O-antigen biosynthesis and leads eventually to a lower amount of LPS I.

Specific aims of this dissertation are as follows:

Aim 1: To investigate the proposed UDP-QuiNAc biosynthetic pathway using a biochemical approach *in vitro*, and synthesize UDP-QuiNAc for use in the study of the initial step of O-antigen biosynthesis. Achievement of this aim is summarized in Chapter Two.

Aim 2: To elucidate the initial step of O-antigen biosynthesis in *R. etli* CE3. Specifically, to investigate the activity of enzymes proposed to be involved in the initial step, to deduce the order in which they exert their activity *in vivo*, and to confirm that QuiNAc is the first sugar of

the O antigen of *R. etli* CE3. Progress toward this aim and interpretation of results are presented in Chapter Three.

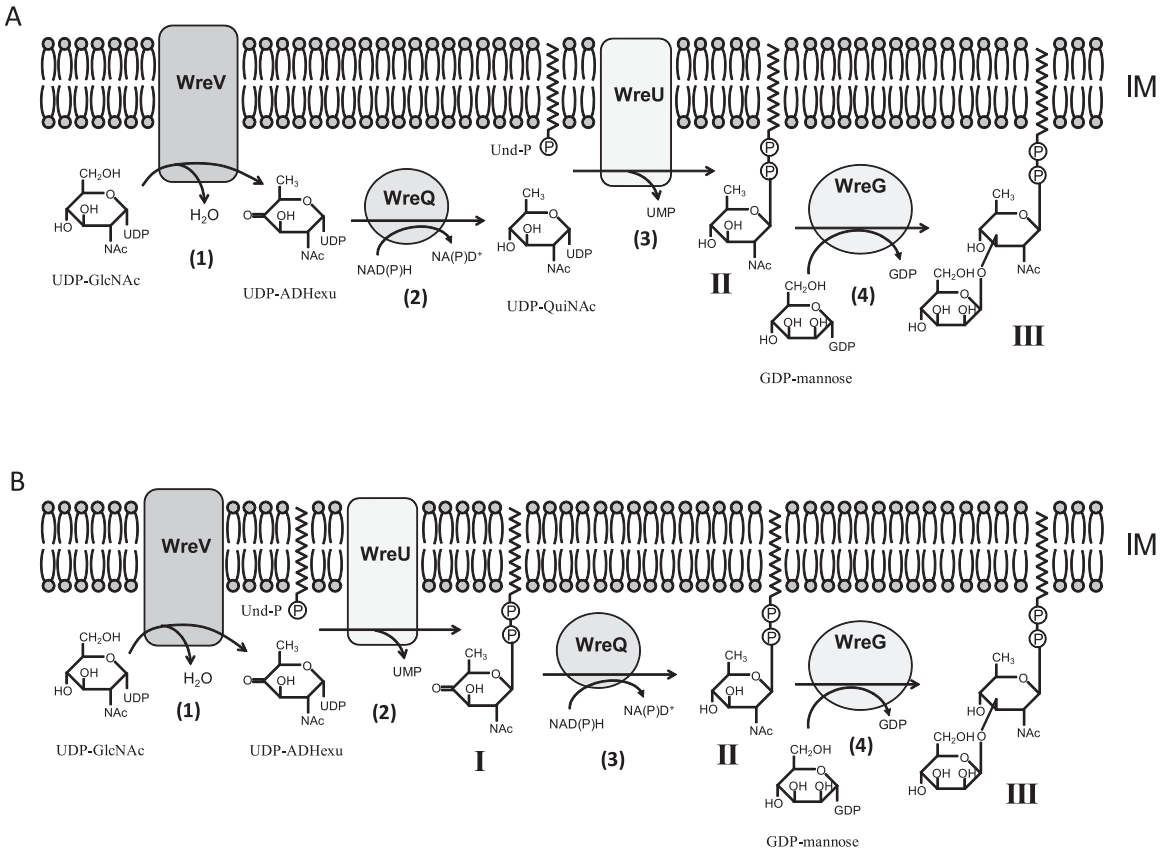


Figure 14. Schematic diagram showing the reactions in each hypothesis. (A) Hypothesis 1. (B) Hypothesis 2. Each hypothesis contains 4 reactions, which are indicated by numbers (1), (2), (3) and (4). The lipidated compounds are represented by roman numbers: I, Und-PP-ADHexu; II, Und-PP-QuiNAc; III, Und-PP-ADHexu. The localization of putative enzymes WreV, WreU, WreQ and WreG are based on computational predictions and similarity to homologous proteins. In this figure, no attempt is made to depict speculated interactions between the proteins.

CHAPTER TWO: IN VITRO BIOSYNTHESIS AND CHEMICAL IDENTIFICATION OF UDP-QUINAC

A. Introduction

Many rare sugars are found in the polysaccharide structures coating bacterial cell surfaces. A moderately rare example is QuiNAc, which so far has been found exclusively in the lipopolysaccharides (LPSs) or capsular polysaccharides of bacteria. Although confined to bacteria, it is found in numerous species, including ones belonging to the genera *Pseudomonas* and *Rhizobium* (4,45).

The QuiNAc residue of LPS in *Rhizobium etli* strain CE3 may play an important role in the biology of this bacterium. An *R. etli* mutant strain (CE166) carrying a Tn*5* transposon in the *wreQ* gene locus (formerly designated *lpsQ*) fails in the infection stage of symbiosis with legume hosts. This mutant has much less O-antigen-containing LPS (LPS I) than normal (only 40% of the wild-type value) and its O antigen lacks QuiNAc (6,35). The QuiNAc is replaced in this mutant by its 4-keto derivative, ADHexu (19). The deficiency in O-antigen amount, but not the absence of QuiNAc, can be suppressed by introducing into this strain recombinant plasmids that share a 7.8-kb stretch of the *R. etli* CE3 *lps* genetic region α, even though this suppressing DNA did not carry the genetic region mutated in strain CE166. Because QuiNAc is still replaced in the resulting strain by ADHexu, the effect of just this abnormality is thereby revealed. The mutant genetically suppressed in this way infects slowly and resulting root nodules are widely dispersed (35). This result suggests that the symbiotic role of LPS may require a specific structural feature conferred by QuiNAc.

QuiNAc has been proposed to be the first sugar of the O antigen of *R. etli* CE3 and the O-antigen repeating unit in *Pseudomonas aeruginosa* O6 (21,43,46). However, these inferences lack evidence from biosynthesis of the polysaccharide *in vitro*. One obstacle to such studies is that the

expected biosynthetic donor of the QuiNAc residue, UDP-QuiNAc, has never been produced *in vitro*.

The proposed biosynthetic pathway of UDP-QuiNAc, based on studies in *R. etli* CE3 and in *P. aeruginosa* O6 (19), involves two steps. In the first step, the precursor UDP-GlcNAc is converted by a 4,6-dehydratase to its 4-keto-6-deoxy derivative, UDP-ADHexu. This reaction has been reproduced *in vitro* by use of several homologous enzymes (37-40,47). It is the common first step in synthesizing many UDP-6-deoxy-sugars from UDP-GlcNAc (36).

In the proposed second step, the 4-keto intermediate UDP-ADHexu is further converted to UDP-QuiNAc by a 4-reductase. This reaction and an enzyme to catalyze it have never been demonstrated *in vitro*. It has been proposed that the *R. etli* CE3 *wreQ* gene encodes the 4-reductase for this step. This gene was previously designated *lpsQ* (19), but all genes involved in *R. etli* CE3 O-antigen biosynthesis have been assigned the symbol *wre* to conform with bacterial polysaccharide genetic nomenclature (21). The predicted amino acid sequence suggests that WreQ belongs to the short chain dehydratase/reductase (SDR) superfamily, characterized by its G*XX*G*XX*G nucleotide-binding motif and the SYK catalytic triad (19,42). As mentioned above, chemical analysis of O antigen from the symbiosis-deficient *wreQ* null mutant showed that QuiNAc was replaced by its 4-keto derivative at the same position (19), which supported that WreQ is the 4-reductase for the last step in QuiNAc synthesis.

P. aeruginosa serotype O6 contains QuiNAc in the O antigen of its B-band LPS. A *wreQ* gene homolog, *wbpV*, was found in its B-band O antigen gene cluster. Disruption of the *wbpV* gene abrogates O-antigen biosynthesis (43). *P. aeruginosa* strains (O6 and PAO1) also contain the *wbpM* gene which encodes a UDP-GlcNAc 4,6-dehydratase (43). The enzymatic activity of WbpM (PAO1) has been biochemically demonstrated *in vitro* (37,39). The *R. etli* CE3 genome contains a *wbpM* gene homolog designated *wreV*. The gene products of both *wbpM* and *wreV* belong to a family of long-chain dehydratases that possess amino-terminal transmembrane domains and the SMK catalytic triad (as opposed to SYK found in the SDR family). A logical

prediction is that these gene pairs (*wreV* and *wreQ* in *R. etli* CE3, *wbpM* and *wbpV* in *P. aeruginosa* O6) encode the enzymes needed to synthesize QuiNAc.

In the work presented in this chapter, the *R. etli* CE3 WreQ protein was purified and its enzymatic activity was investigated *in vitro*. The nucleotide-sugar product of the WreQ-catalyzed reaction was chemically identified to be UDP-QuiNAc. The results confirmed that WreQ can act as a 4-reductase and provide the first demonstration of UDP-QuiNAc synthesis *in vitro*.

B. Site-directed mutagenesis of *R. etli wreV* and *wreQ* gene

In both of the hypotheses presented in Chapter One· L, QuiNAc synthesis begins with the 4,6-dehydration of the precursor UDP-GlcNAc, which is proposed to be catalyzed by WreV in *R. etli* CE3. Disruption of the *wreV* gene is expected to abolish QuiNAc synthesis. Consistent with this prediction, a previous isolated mutant (strain CE168) carrying the transposon Tn*5* in the *wreV* gene locus completely lacked LPS I (6). The *wreQ*::Tn*5* mutant CE166 has been thoroughly investigated in previous studies (19,35) and its phenotype was presented in Chapter One· J. Because Tn5 insertions tend to have polar effects (48,49), *wreV* and *wreQ* genes were mutated separately by insertion of nonpolar antibiotic-resistance cassettes in this study.

Mutagenesis of wreV *gene*— the *wreV* gene was disrupted by deleting the sequence between two internal PstI restriction sites and inserting either a gentamicin resistance cassette (Gm) or a kanamycin resistance cassette (Km). This approach has been shown to not cause polar effects on genes downstream in an operon (50). The resulting mutant strains CE568 (*wreV*::Gm) and CE569 (*wreV*::Km) showed a lack of LPS I (LPS I^-) phenotype (Fig. 15A, lane 3 and 4). Both CE568 and CE569 can be complemented by introducing the wildtype *wreV* gene on the pFAJ1708 vector (51), as shown by the reappearance of the LPS I band (Fig. 15B, lane 3 and 5).

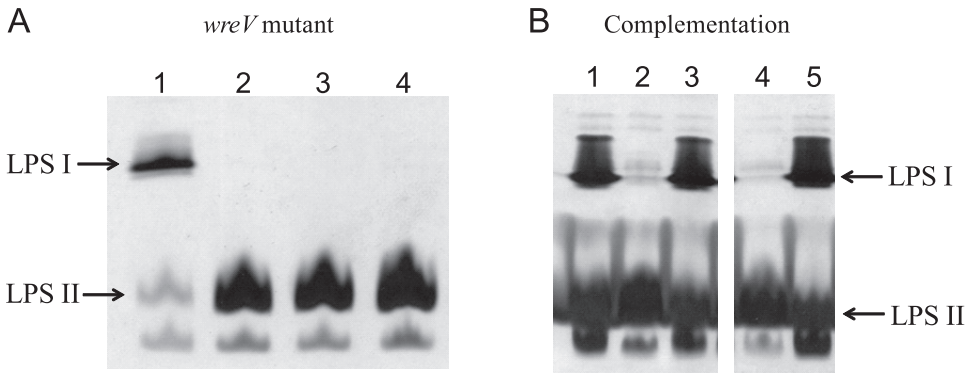


Figure 15. SDS-PAGE analysis of the LPS phenotype of the *wreV* mutant strains and complemented strains. (A) *R. etli wreV* mutants are devoid of LPS I. Lanes: lane1, CE3 (wildtype); lane 2, CE168 (*wreV*::Tn*5*); lane 3, CE568 (*wreV*::Gm); lane 4, CE569 (*wreV*::Km). (B) *R. etli wreV* mutants can be complemented by wildtype *wreV* gene. Lanes: lane 1, CE3; lane 2, CE568; lane 3, CE568/pTL7; lane 4, CE569; lane 5, CE569/pTL7. pTL7 is a pFAJ1708 vector carrying the wildtype *wreV* gene.

Mutagenesis of wreQ *gene*— the *wreQ* gene was disrupted by the same deletion-insertion approach as described above. The *wreQ*::Km mutant CE572 was obtained but no *wreQ*::Gm mutant could be isolated. Strain CE572 showed the same LPS phenotype—reduced LPS I amount, as the *wreQ*::Tn*5* mutant CE166 when analyzed by SDS-PAGE (Fig. 16A, lane 3), and it can be complemented by the wildtype *wreQ* gene (Fig. 16B, lane 3). This result confirmed that the reduced LPS I amount is caused by the mutation in the *wreQ* gene. However, since the LPS composition of strain CE572 was not investigated, CE166 (*wreQ*::Tn*5*) will be used as the *wreQ* mutant strain in the following experiments in this dissertation.

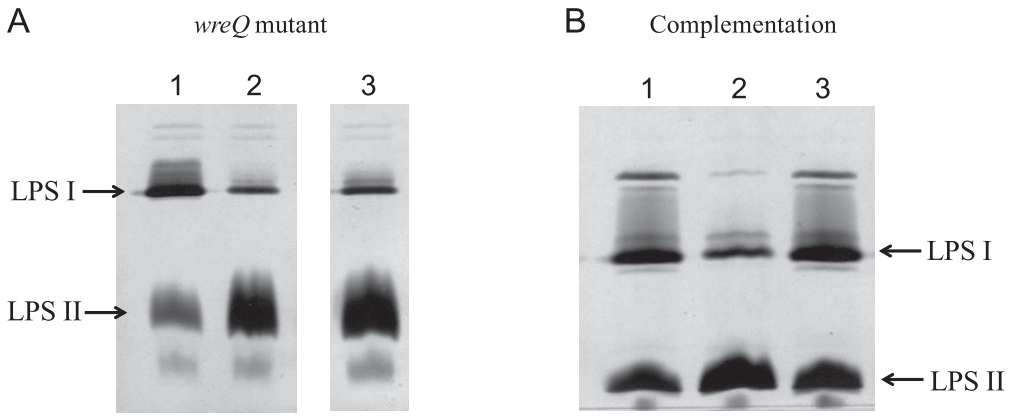


Figure 16. SDS-PAGE analysis of the LPS phenotype of the *wreQ* mutant strains and complemented strains. (A) *R. etli wreQ* mutants show a reduced LPS I amount compare to the wildtype. Lanes: lane 1, CE3 (wildtype); lane 2, CE166 (*wreQ*::Tn*5*); lane 3, CE572 (*wreQ*::Km). (B) Strain CE572 complemented with wildtype *wreQ* gene. Lanes: lane 1, CE3; lane 2, CE572; lane 3, CE572/pJBQ3. pJBQ3 is a pFAJ1708 vector carrying the wildtype *wreQ* gene (19).

C. Cloning, Expression and Purification of WreQ

The *wreQ* gene was amplified from chromosomal DNA of *R. etli* CE3 and cloned into the pET15b vector. The resultant plasmid encodes a WreQ protein with an amino terminal six-histidine (His_6) tag. Following overexpression, this His_6-WreQ protein was purified from the soluble fraction by nickel-affinity chromatography (Fig. 17A).

The modified *wreQ* gene (*His_6-wreQ*) was determined to be functional *in vivo* due to its ability to complement the *R. etli wreQ*::Tn*5* mutant strain CE166. In the complemented mutant, the level of O-antigen-containing LPS (LPS I) was restored to normal (Fig. 18A), and the presence of quinovosamine (derived from QuiNAc) in the LPS was confirmed by sugar composition analysis (Fig. 18B). The WreQ homolog WbpV from *P. aeruginosa* O6 also complemented the same mutant (Fig. 18A, C). All of these results match those reported previously (19) when the wild-type *wreQ* gene was used to complement this same mutant.

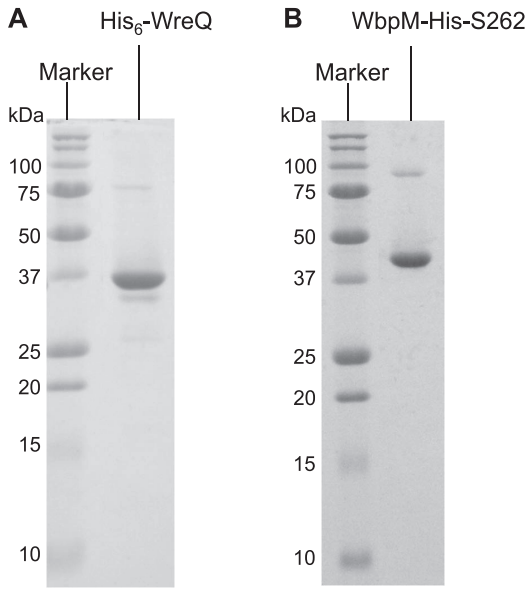


Figure 17. Commassie blue-stained SDS-PAGE (15% gel) showing purified proteins used for *in vitro* biosynthesis of UDP-QuiNAc. (A) His_6-WreQ. (B) WbpM-His-S262. The apparent size of the purified His_6-WreQ protein agreed with its predicted size of 35.3 kDa. The apparent size of the purified WbpM-His-S262 protein is consistent with the published result (37) (46.1 kDa).

D. Generation of WreQ substrate—UDP-ADHexu

Attempts to express the putative UDP-GlcNAc 4,6-dehydratase WreV in *E. coli* for purification and characterization of its activity have been unsuccessful. To generate the putative substrate UDP-ADHexu for investigating the enzymatic activity of WreQ, a soluble truncated version of the WbpM protein, His-S262 (37), was chosen to catalyze the 4,6-dehydration of UDP-GlcNAc. The conversion of tritium (^{3}H)-labeled UDP-GlcNAc to a new compound upon the addition of purified WbpM-His-S262 protein (Fig. 17B) was shown by thin layer chromatography (TLC)-autoradiography (Fig. 20, lane 1 and 2). The same result was obtained from a nonradioactive reaction as shown by TLC with thymol staining (Fig. 19, lane 1 and 2). The new compound was purified from unstained TLC plates (Fig. 19, lane 3). Its 1D and 2D ^{1}H-NMR spectra (discussed below) agreed with the results of previous studies on the WbpM homolog Cj1120c (PglF) (39). The evidence that this product is the UDP-ADHexu will be presented later in this chapter.

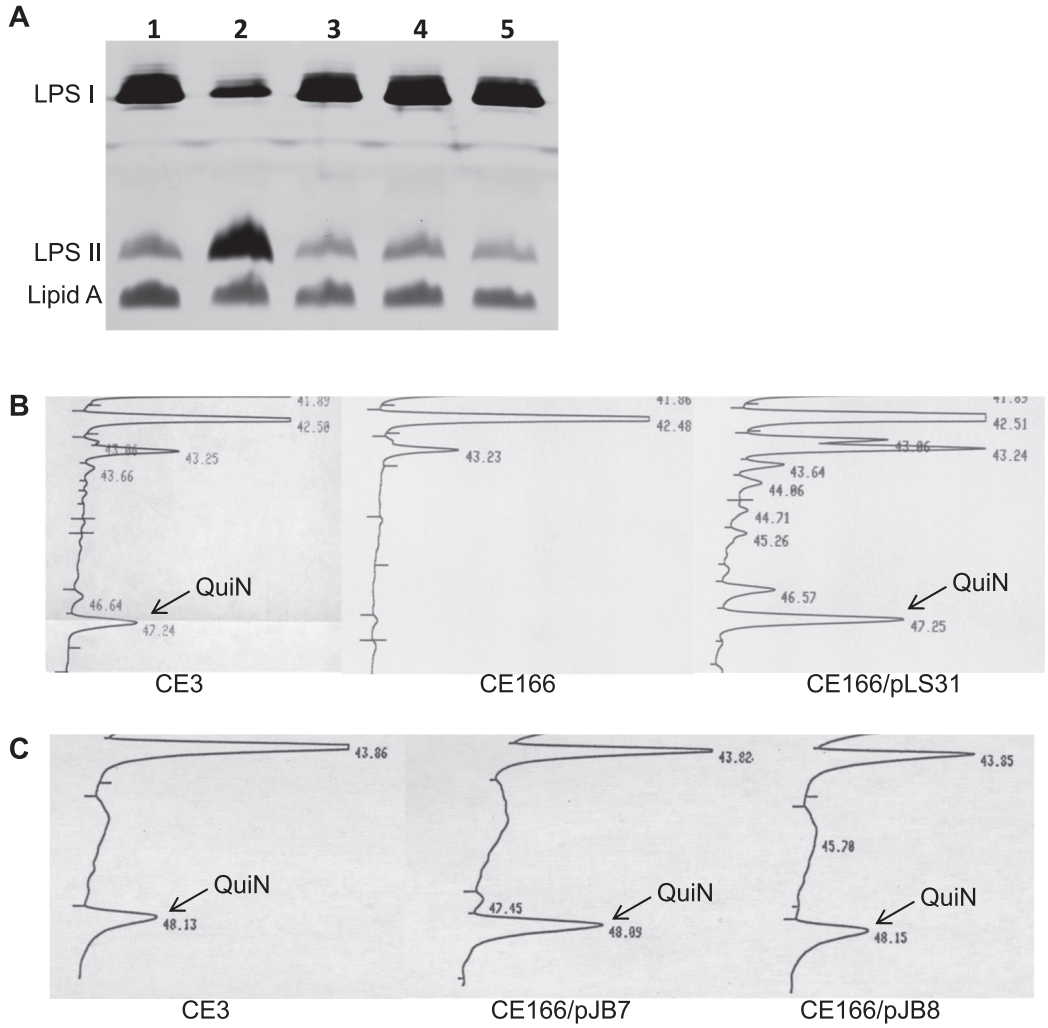


Figure 18. Complementation of the *Rhizobium etli wreQ*::Tn*5* mutant strain, CE166. (A) Silver-stained SDS-PAGE (18% gel) of LPS samples from whole-cell lysates. Lanes: lane 1, CE3; lane 2, CE166; lane 3, CE166/pLS31; lane 4, CE166/pJB7; lane 5, CE166/pJB8. LPS I, O-antigen-containing LPS; LPS II, core oligosaccharide-Lipid A lacking O antigen. (B, C) Relative content of QuiNAc among the sugars of crude LPS as revealed by GC analysis after conversion of sugars to alditol acetates. CE3, wild-type *R. etli* strain; CE166, *R. etli* strain with a Tn*5* insertion in *wreQ*; CE166/pLS31, CE166 strain carrying pLS31 vector which expresses His_6-WreQ; CE166/pJB7 and CE166/pJB8, CE166 strain carrying pJB7 or pJB8 plasmids, each carrying the *wbpV* gene. QuiN, quinovosamine.

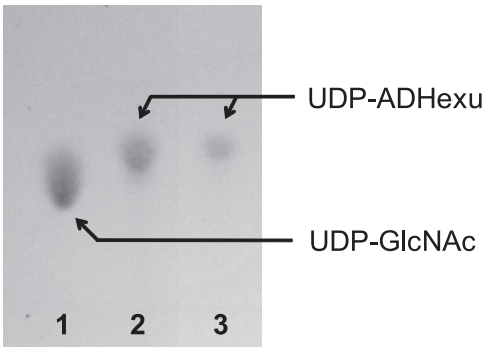


Figure 19. Thymol-stained TLC plate showing the purification of UDP-ADHexu. Lanes: lane 1, UDP-GlcNAc, 25 nmol; lane 2, WbpM-reaction sample, 10 nmol total nucleotide sugar; lane 3, purified UDP-ADHexu candidate, estimated amount 5 nmol.

E. Substrate conversion by WreQ

A two-step reaction was carried out to test the enzymatic activity of WreQ. Following the first-step reaction catalyzed by WbpM, purified His_6-WreQ protein and NAD(P)H were added to initiate the second-step reaction. A new compound later shown to be UDP-QuiNAc was produced (Fig. 20, lane 3 and 4). Both the WreQ enzyme and NAD(P)H were required to produce this new compound (Fig. 20, lane 6 and 7). WreQ did not use UDP-GlcNAc as substrate; i.e., no conversion occurred in the absence of WbpM (Fig. 20, lane 5).

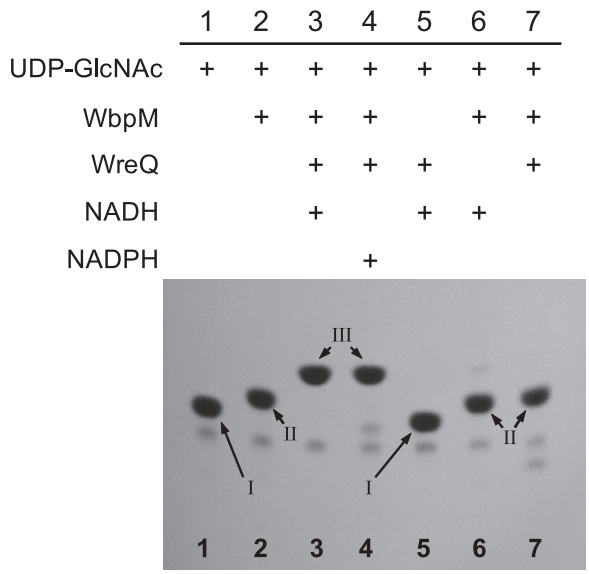


Figure 20. Autoradiograms of TLC-separated WbpM-WreQ reaction and control reactions. Lanes: lane 1, no enzymes were added to the nucleotide sugar substrate UDP-GlcNAc; lane 2, WbpM reaction for synthesizing UDP-ADHexu; lane 3, WbpM-WreQ reaction, for synthesizing UDP-QuiNAc; lane 4, NADPH in place of NADH; lane 5, negative control, no WbpM; lane 6, negative control, no WreQ; lane 7, no NAD(P)H. NMR analysis confirmed the identities of the spots as follows: Spot I, UDP-GlcNAc; Spot II, UDP-ADHexu; Spot III, UDP-QuiNAc.

F. GC-MS analysis of the WbpM-WreQ reaction product

To gain initial evidence that UDP-QuiNAc was synthesized in the WbpM-WreQ coupled reaction, the reaction mixture was treated at 100 °C with acid to release the sugars from the UDP linkages. The resulting compounds were reduced and acetylated to convert sugars into deuterated alditol acetates, followed by combined gas chromatography-mass spectrometry (GC-MS) analysis (Fig. 21A, B). Pure UDP-GlcNAc was analyzed after the same treatment as a control (Fig. 21C, D).

When the WbpM-WreQ reaction sample was analyzed, a major peak eluted at 26.61 min (Fig. 21A). It fragmented into ions with m/z ratios 85, 103, 145 which identified it as a 2-amino sugar, and ions with m/z ratios 201, 260, 302 (Fig. 21B), which were 58 mass units smaller than the corresponding ions of the alditol acetate derived from GlcNAc (Fig. 21D). This difference agreed exactly with predicted differences at the C6 position, where the GlcNAc derivative is

CH_2OCOCH_3 (molecular mass=73) and QuiNAc derivative is CH_3 (molecular mass=15). This data supported the existence of the QuiNAc moiety in the reaction mixture, as expected if UDP-QuiNAc is produced.

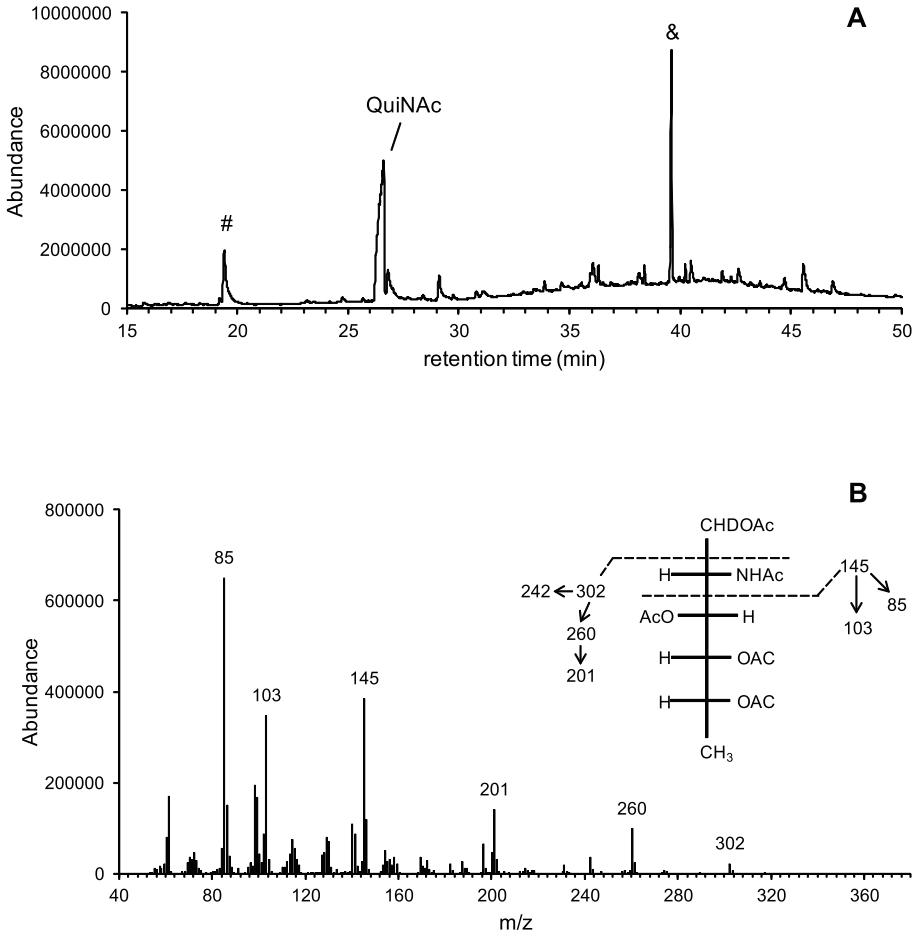


Figure 21. (A, B)

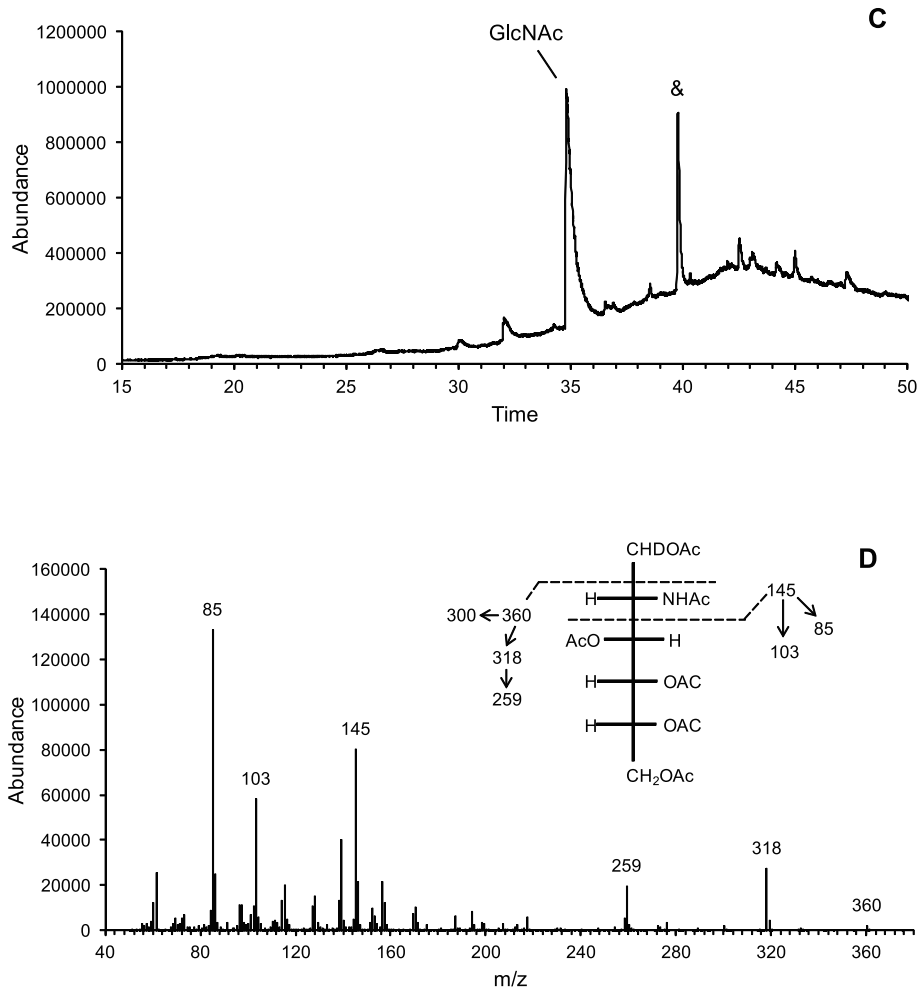


Figure 21. GC-MS analysis of chemical derivatives after converting sugars to alditol acetates. (A) GC profile of the WbpM-WreQ reaction sample after alditol acetate treatment. (B) MS spectrum of the QuiNAc derivative peak in panel A. *inset*, predicted MS fragmentation patterns of the alditol acetates derived from QuiNAc. (C) GC profile of UDP-GlcNAc after alditol acetate treatment. (D) MS spectrum of the GlcNAc derivative peak in panel C. *inset*, predicted MS fragmentation patterns of the alditol acetates derived from GlcNAc. MS peaks corresponding to the predicted fragmentation pattern were shown with numbers above. The peak designated with "&" in panel A and C is a common non-sugar peak. The peak designated with "#" in panel A may have resulted from other components in the enzyme reaction as it was not detected in panel C.

G. NMR analysis of the WreQ-mediated production of UDP-QuiNAc

Compounds in a non-radioactive WbpM-WreQ reaction mixture were separated by TLC, and the UDP-QuiNAc candidate was purified from an unstained TLC plate (Fig. 22). The purified compound was analyzed by one and two-dimensional NMR spectroscopy (Fig. 23 and Fig. 24).

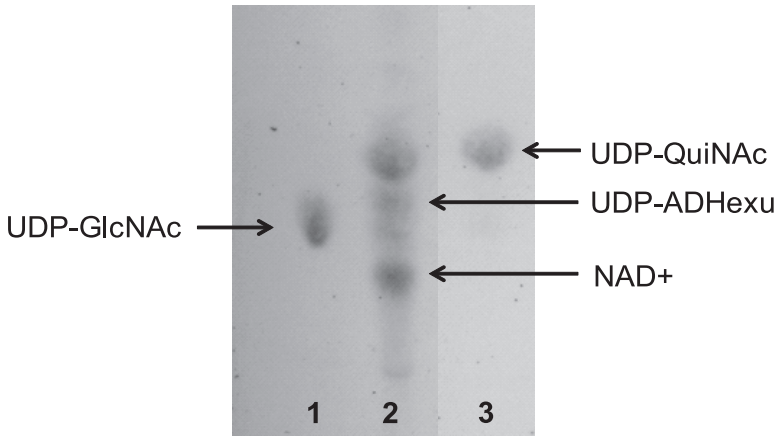


Figure 22. Thymol stained TLC plate showing the purification of UDP-QuiNAc. Lanes: lane 1, UDP-GlcNAc, 20 nmol; lane 2, WbpM-WreQ reaction sample, 20 nmol total nucleotide sugar; lane 3, purified UDP-QuiNAc candidate, estimated amount 7.5 nmol.

Assignment of proton resonances of the WreQ reaction product utilized a combination of NOESY, TOCSY and COSY correlations. The 2D ^{1}H-^{1}H TOCSY experiment allowed establishing the resonance frequencies of all protons in the QuiNAc sugar ring (Fig. 23B). The UDP ribose ring protons were observed between 4.174 and 4.344 ppm and at 5.959 ppm. The one at 5.959 ppm was correlated to the uracil base proton at 7.935 ppm (not shown). The UDP resonances did not significantly change from UDP-GlcNAc to UDP-ADHexu and to UDP-QuiNAc; they will not be discussed. Contaminants uncoupled from UDP-QuiNAc resonances were observed in small amounts in the UDP-QuiNAc sample as isolated spin systems at low intensity between 3.624 ppm and 3.814 ppm; at 1.677 and 2.144 ppm; and at 1.221 ppm (Fig. 23B). The hexose protons assigned to the UDP-QuiNAc resonance chemical shifts (Table 1) are explained below.

Table 1. Summary of the proton chemical shifts and multiplet structures for the hexose ring of UDP-QuiNAc and UDP-ADHexu at 30 °C. Standard deviation of the chemical shift values is on the order of 0.004 ppm (as reported by Sparky). Accuracy of J-coupling constant estimates is ±0.2 for for UDP-D-QuiNAc and ±0.6 Hz for UDP-2-acetamido-2,6-dideoxy-D-xylo-4-hexulose.

Compound	Proton	Chemical shift (ppm)	Multiplet structure	J-coupling constants (Hz)
UDP-QuiNAc	H-1	5.422	1:1:1:1 quadruplet	$^3J_{H1,H2}$ = 3.6 $^3J_{H1,P}$ = 7.2
	H-2	3.978	unresolved overlapped multiplet	
	H-3	3.723	1:2:1 triplet	$^3J_{H3,H4}$ = 9.8 $^3J_{H2,H3}$ = 9.8
	H-4	3.231	1:2:1 triplet	$^3J_{H3,H4}$ = 9.3 $^3J_{H4,H5}$ = 9.3
	H-5	3.977	unresolved overlapped multiplet	
	H-6	1.271	1:1 doublet	$^3J_{H6,H5}$ = 6.0
UDP-ADHexu[a]	H-1	5.433	poorly resolved multiplet	
	H-2	4.091	unresolved overlapped multiplet	$^3J_{H3,H2}$ = 10.8
	H-3	3.803	doublet	
	H-5	4.091	unresolved overlapped multiplet	
	H-6	1.218	doublet	$^3J_{H6,H5}$ = 5.4

[a] Due to a lower concentration of the latter sample the spectral quality did not allow confident measurement of the smaller ^{3}J-coupling constants at H1 position.

The methyl proton doublet centered at 1.271ppm was identified as H6 by comparison with the proton spectrum of UDP-GlcNAc, which shows the only methyl resonance at 2.044 ppm originating from its acetamido group. Inspection of the 2D ^{1}H-^{1}H COSY revealed that the neighbor of H6, the H5 proton, resonates at 3.977 ppm (Fig. 23C). The peak of H5 had significant asymmetry indicating overlapping resonances, and showed four COSY cross-peaks.

The proton resonance at 5.442 ppm was assigned to be H1 by similarity to other sugar nucleotides (52). Inspection of the 2D ^{1}H-^{1}H NOESY correlation map (Fig. 23D) reveals a strong (and the only) NOE cross-peak of H1 at 5.442 and 3.977 ppm indicating that H5 must be overlapping with H2 at this chemical shift because H5 is too far away to produce NOE to H1. (In the COSY data the H1 resonance was mirrored into the 4.0 ppm region, thus obscuring the expected cross-peak with H2).

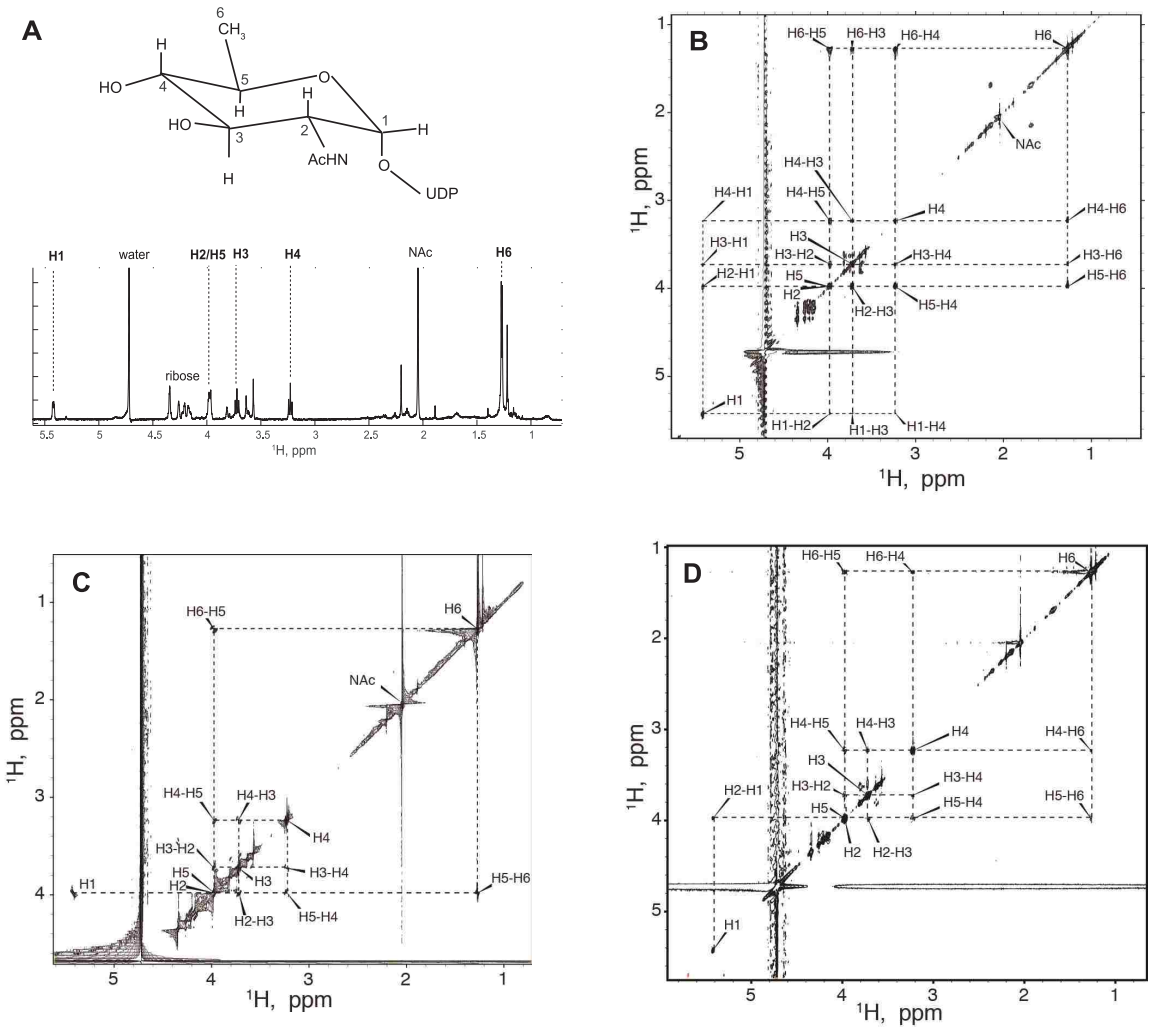


Figure 23. Structural determination for UDP-QuiNAc. (A) Structure and proton NMR resonance assignments of the reaction product deduced to be UDP-QuiNAc. (B-D), two dimensional ^{1}H-^{1}H TOCSY, COSY, NOESY analysis of UDP-QuiNAc: (B), 2D ^{1}H-^{1}H TOCSY correlation map for UDP-QuiNAc. (C) 2D ^{1}H-^{1}H COSY correlation map for UDP-QuiNAc. The experiment was acquired in real-time mode; therefore, all resonances with chemical shifts beyond water appear mirrored. (D) 2D ^{1}H-^{1}H NOESY correlation map for UDP-QuiNAc.

Out of the two remaining resonances at 3.723 and 3.231 ppm, we assigned H4 to 3.231 according to the following considerations: (1) the TOCSY cross-peak of H6-H3 is weaker comparatively to the H6-H4 cross-peak, in accord with the greater number of bonds separating H6 and H3 spins (Fig. 24A, top panel); (2) an opposite relationship is observed for TOCSY cross-peaks of H1-H3 and H1-H4—the latter being the weakest of the two, corresponding to the greater

number of intervening bonds (Fig. 24A, bottom panel); (3) this resonance is absent in the 4-keto intermediate, which would not have a proton at position 4 (see Fig. 24B).

The configuration of the H4 and H5 in UDP-QuiNAc was determined to be axial based on analysis of ^{3}J-coupling constants. Both H3 and H4 are symmetrical triplets indicating identical ^{3}J-coupling constants of 9-10 Hz for all H2-H3, H3-H4, and H4-H5 pairs (Fig. 24B, top trace). The large value of the ^{3}J-coupling constants is indicative of the axial orientation of coupled protons (53); therefore, we conclude that H4 and H5 are axial, which leads to the structure shown in the Fig. 23A, corresponding to the configuration of a QuiNAc ring.

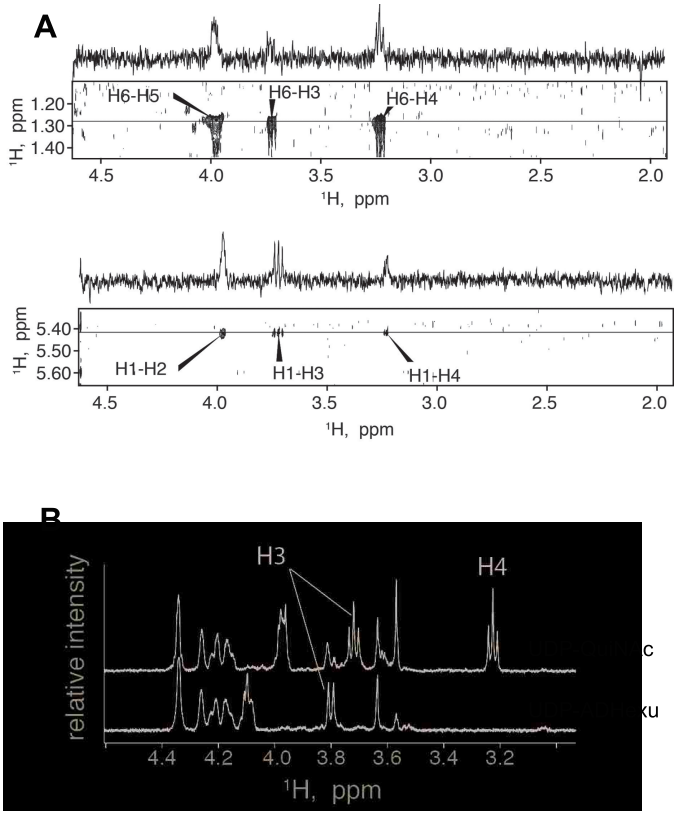


Figure 24. Assignment of proton chemical shifts for H3 and H4. (A) Strips of 2D ^{1}H-^{1}H TOCSY correlation map for UDP-QuiNAc and one-dimensional slices taken at frequencies of H6 (*top panel*) and H1 (*bottom panel*) in indirect dimension demonstrating relative intensities of the cross-peaks. (B) Comparison of one-dimensional proton NMR spectra of purified UDP-QuiNAc (*top trace*) and UDP-ADHexu (*bottom trace*).

H. NMR analysis of the WreQ substrate for UDP-QuiNAc synthesis—UDP-ADHexu

To confirm that the substrate in the WreQ-catalyzed reaction (second step) is truly UDP-ADHexu, the product of the WbpM reaction (first-step) was purified and analyzed by NMR. Assignment of NMR signals was performed in a similar fashion as for the UDP-QuiNAc. The TOCSY correlation map demonstrated that the WbpM reaction product contains two isolated proton spin systems with signals of one assigned to H6 and H5 while the second spin system included H1, H2, and H3. The H2 and H5 resonances overlap in the same fashion as in the TOCSY spectra of UDP-QuiNAc (Fig. 25A), but the absence of a H6-H3 cross peak indicates a break in the continuous path for the TOCSY transfer from H6 to H3 due to the H4 proton missing in the 4-keto intermediate.

Panel B of Fig. 24 compares the one-dimensional spectra of UDP-QuiNAc and the WbpM reaction product demonstrating the absence of the H4 signal in the 4-keto intermediate spectrum (bottom trace). Additionally, the H3 signal becomes a doublet indicating that H3 has only one ^{3}J-coupling partner in this molecule, which must be H2. These findings confirm that the WbpM reaction product and substrate of WreQ was, indeed, UDP-ADHexu, with the structure and the proton assignment shown in Fig. 25, panel B.

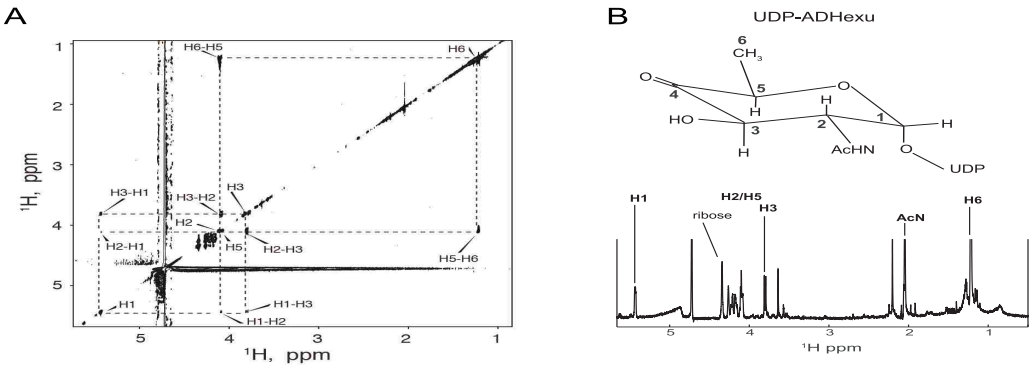


Figure 25. Structural determination for UDP-ADHexu. (A) TOCSY correlation map obtained using a purified sample of the WbpM reaction product. (B) Molecular structure and proton NMR resonance assignments of UDP-ADHexu.

I. Direct observation of the reaction progress by NMR

Stability of the compounds in the NMR tube was directly assessed by comparison of one-dimensional proton spectra acquired before and after two-dimensional experiments. Samples were found to be completely stable over the entire time of data acquisition.

To gain information about reaction kinetics, we performed the enzyme incubation in an NMR spectrometer and directly monitored the WbpM-WreQ reaction progress. Initial accumulation of the UDP-ADHexu and its subsequent conversion into UDP-QuiNAc was monitored by ^{1}H NMR spectroscopy at 30 $^\circ$C. Fig. 26 shows an overlay of one-dimensional proton spectra taken during the reaction course. The initial reaction mixture (black trace) contained UDP-GlcNAc, NADH, and purified WreQ, but no WbpM. Addition of WbpM led to quick accumulation of UDP-ADHexu (reflecting 4,6-dehydratase activity of WbpM) as indicated by appearance of the new H6 methyl proton doublet centered at 1.217 ppm. Activity of WreQ led to subsequent conversion of UDP-ADHexu to UDP-QuiNAc that was detected by the appearance of its H6 methyl doublet centered at 1.271 ppm (and decreasing H6 methyl signal of UDP-

ADHexu). Because the predicted product UDP-QuiNAc retains the methyl group in position 6, the chemical shift change must reflect the modification of the hexose ring at position 4. The WbpM-catalyzed reaction was significantly faster than that catalyzed by WreQ. Most UDP-ADHexu was produced within the first 10 minutes (black to blue traces), while subsequent slow conversion to UDP-QuiNAc continued during overnight incubation and, even then, was not complete (blue to red traces).

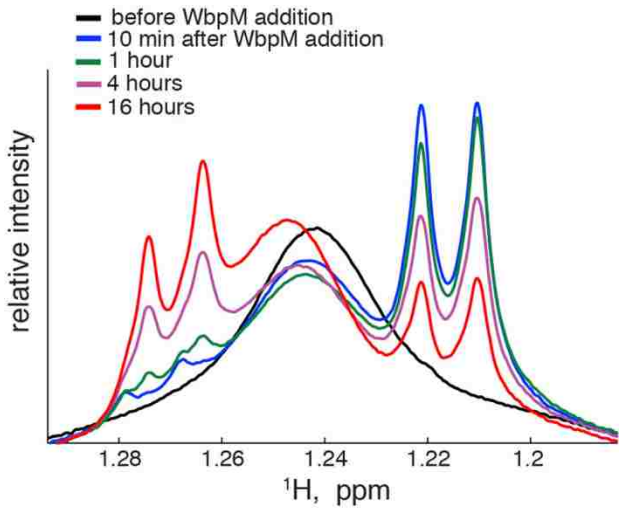


Figure 26. One-dimensional proton NMR spectra obtained from the reaction mixture at different incubation times. A broad resonance at 1.241 ppm is due to methyl groups of the detergent Triton X-100 present in the reaction mixture. This reaction was performed with 1 mM UDP-GlcNAc, 1 mM NADH, 48 µg WbpM and 54 µg WreQ in a total volume of 300 µl. WbpM was added last to start the reaction.

J. DISCUSSION

Quinovosamine was first identified in bacterial polysaccharides more than 50 years ago (54,55). Attempts to establish the pathway of its biosynthesis foundered on the use of crude extracts and the problem of demonstrating a specific reaction by adding NADH to a crude extract (56). Only now, with the use of a genetic approach leading to a facile enzyme purification, has an enzymatic pathway for its synthesis been demonstrated *in vitro*. In this report, we present

evidence that UDP-QuiNAc can be synthesized *in vitro* from UDP-GlcNAc via the two steps proposed previously (19) (Fig. 10, Fig. 27).

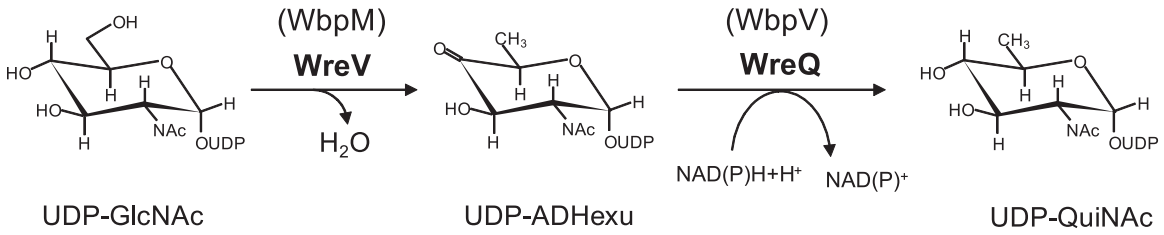


Figure 27. The proposed biosynthesis pathway of UDP-QuiNAc in *Rhizobium etli* CE3 and *Pseudomonas aeruginosa* O6. The *P. aeruginosa* O6 proteins are indicated in parentheses.

The first step is catalyzed by UDP-GlcNAc 4,6-dehydratase, resulting in the 4-keto intermediate UDP-ADHexu. A soluble truncated version of the WbpM protein from *P. aeruginosa* was used in this study to catalyze this reaction *in vitro* (37). The *R. etli* CE3 *wreV* gene encodes a protein that is homologous to WbpM. An insertion mutation in *wreV* abolished O-antigen biosynthesis in *R. etli*, and *wbpM* complemented this mutant (Fig. 28). Hence, WreV is inferred to be the 4,6-dehydratase responsible in *R. etli* for the first step of UDP-QuiNAc synthesis.

The second step in Fig. 27 is catalyzed by a 4-reductase. Previous studies (19,35) suggested that *wreQ* encodes this enzyme in *R. etli* CE3. The *in vitro* reaction in this study confirmed this prediction. WreQ stereospecifically reduces the 4-keto intermediate to UDP-QuiNAc, instead of its C-4 epimer UDP-*N*-acetyl-D-fucosamine (UDP-FucNAc). A homolog of WreQ in *P. aeruginosa* O6, encoded by *wbpV*, has been proposed to catalyze the same reaction (43). As predicted, the *wbpV* gene can complement the *R. etli wreQ*::Tn*5* mutant (Fig. 18, A, C).

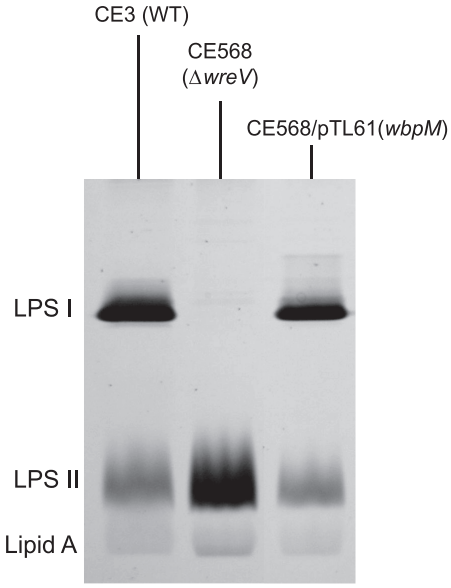


Figure 28. *Pseudomonas aeruginosa wbpM* gene can complement *Rhizobium etli wreV* mutant (CE568, *ΔwreV*). Silver-stained SDS-PAGE (18% gel) of LPS samples from whole-cell lysates. LPS I, O-antigen-containing LPS; LPS II, core oligosaccharide-Lipid A lacking O antigen. CE3, wild-type *R. etli* strain; CE568, *R. etli* strain with a gentamicin resistance cassette insertion in *wreV* (Δ*wreV*); CE568/pTL61, CE568 strain carrying plasmid pTL61 which expresses the full-length WbpM.

The second reaction catalyzed by WreQ in Fig. 27 was relatively slow. The low activity may be an artifact of one or more features of the *in vitro* reaction. The His_6-WreQ protein used *in vitro* has additional amino acids at the amino-terminus, which could be affecting the structure at the active site and slowing catalysis. Another consideration is that, during dialysis to reduce the concentration of imidazole used in purification of the His-tagged protein, the His_6-WreQ protein precipitated in the dialysis tubing. It was found that adding detergent Triton X-100 alleviated the precipitation, but the presence of detergent (0.1% of reaction volume) may have affected the reaction rate. It is unclear why His_6-WreQ precipitated during dialysis. One possible explanation is that WreQ normally associates with the membrane or other proteins in *R. etli* CE3 and that they are required for its stability. If that is the case, overproduction of WreQ in *E. coli*, which is devoid of its naturally-associating partner, may render it unstable.

The pathway intermediate UDP-ADHexu was relatively stable. Although other 4-keto, 6-deoxy sugars have been reported to be unstable (52,57-59), this 4-keto-6-deoxy derivative of

UDP-GlcNAc was able to withstand purification by TLC, and was stable for several days during NMR data acquisition. Thus, we were able to obtain both this intermediate and UDP-QuiNAc in stable, relatively pure form.

Analysis of O-antigen mutants and the sequences of the *wre* genes for O antigen synthesis have prompted speculation that O-antigen biosynthesis in *R. etli* CE3 begins with reactions that result in 1-pyrophospho-QuiNAc attached to a lipid carrier (21). As noted in the introduction, a previous study showed that QuiNAc was replaced by its 4-keto derivative in the LPS of the *wreQ*::Tn*5* mutant strain, CE166 (19). This observation leads to questions regarding the substrate specificities of WreQ and the putative GTase that catalyzes the initial sugar-lipid linkage. It is conceivable that the initiating GTase works better with UDP-ADHexu as substrate. If so, the resulting lipidated ADHexu may be a better substrate for WreQ than UDP-ADHexu. This also could explain the relatively slow catalysis of WreQ noted in this study. The compounds generated *in vitro* in this study provide an essential starting point for testing the speculations.

CHAPTER THREE: THE INITIAL STEP OF O-ANTIGEN BIOSYNTHESIS IN *RHIZOBIUM ETLI* CE3

A. Introduction

Despite tremendous structural variations, the biosynthesis of all characterized O antigens is believed to begin with a conserved initial step, which involves the formation of an Und-PP-linked glycose by transfer of the sugar-1-P residue from an NDP-sugar to the lipid carrier Und-P. This step is catalyzed by an initiating GTase which differs from other GTases involved in O-antigen biosynthesis in that it recognizes a hydrophobic lipid carrier rather than a sugar acceptor, and that it transfers the sugar-1-P residue rather than the sugar alone.

The first O-antigen sugar in *Rhizobium etli* CE3 has been proposed to be QuiNAc, based on the following evidence. First of all, structural studies indicated that QuiNAc was linked to the non-reducing end of a Kdo (3-deoxy-D-manno-oct-2-ulosonic acid) residue (4,31) and biosynthetic/genetic studies strongly supported that this Kdo is part of the biosynthetic core of the LPS (5,32-34). Secondly, LPS composition analysis of O-antigen mutants that either lack O antigen completely or produce truncated O antigens that are in much lower amount (than the full-length O antigen produced in WT) revealed that QuiNAc is always significantly less in amount than core-specific sugars such as galactose. This strongly suggested that QuiNAc is not a core sugar (12). Lastly, the first sugar of many O antigens is an amino sugar. Such is the case for most *E. coli* strains in which GlcNAc is the first sugar (26), and was recently demonstrated for *Aeromonas hydrophila* AH-3 in which O antigen is initiated with GalNAc (60). In addition, FucNAc and QuiNAc have been proposed to initiate the O repeating unit in *Pseudomonas aeruginosa* serotype O5 and O6, respectively (43,61).

Based on mutational analysis, the *wreU* gene located on plasmid B is predicted to encode the initiating GTase for O-antigen biosynthesis in *R. etli* CE3 (12). The WreU protein is homologous to the C-terminal domain of WbaP ($WbaP_{CT}$), which was shown to be a UDP-

Gal:Und-P Gal-1-P transferase (29). Furthermore, the *wreU*::Km mutant strain (CE566) lacks the LPS I band when analyzed by SDS-PAGE. Sugar composition analysis of its LPS revealed the lack all O-antigen-specific sugars including QuiNAc (12).

It is most straightforward to propose that WreU initiates O-antigen biosynthesis in *R. etli* CE3 by transferring QuiNAc-1-P from UDP-QuiNAc to Und-P. However, the LPS phenotype of *wreQ*::Tn*5* mutant strain CE166 suggests that WreU is able to transfer ADHexu-1-P to Und-P, possibly from UDP-ADHexu, the intermediate in the UDP-QuiNAc synthesis. Thus, both UDP-QuiNAc and UDP-ADHexu will be tested as potential substrates of WreU.

R. etli WreQ was used in the form of purified enzyme *in vitro* towards synthesizing UDP-QuiNAc, as described in Chapter Two. Although WreQ was able to catalyze the reduction of UDP-ADHexu to UDP-QuiNAc, the reaction it catalyzed was relatively slow. It raises the question of whether UDP-ADHexu is the natural substrate of WreQ *in vivo*. Considering that the initiating GTase (putatively WreU) may use UDP-ADHexu as substrate and transfer ADHexu-1-P to Und-P, a possible alternative is that the true substrate of WreQ *in vivo* is Und-PP-ADHexu and WreQ reduces it to Und-PP-QuiNAc. In this scenario, it is very likely that UDP-QuiNAc has never been made in the cell.

According to the structure of *R. etli* CE3 O antigen (4), if QuiNAc is the first O-antigen sugar, mannose is the second. The GTase for adding this Man residue was proposed to be encoded by the *wreG* gene, located within the chromosomal O-antigen genetic cluster. Several lines of evidence support this role of WreG (12): the primary sequence of WreG is highly homologous to the mannosyltransferase WbdC of *E. coli* O8 and O9a; the *wreG* mutant lacks the wild-type LPS I band and has no visible truncated LPS I band that can be resolved from LPS II on SDS-PAGE gels, suggesting that its O antigen is truncated at an early step; LPS sugar composition revealed that the *wreG* mutant lacks all other O-antigen specific sugars but contains QuiNAc (12). It is most straightforward to postulate that WreG normally adds mannose to the Und-PP-QuiNAc acceptor. However, WreG should be able to add mannose to the Und-PP-

ADHexu acceptor as this reaction apparently occurs in CE166. Futhermore, a variation of hypothesis 2 is that WreG normally adds mannose to Und-PP-ADHexu before the reduction ADHexu to QuiNAc by WreQ. This possibility cannot be ruled out until further evidence is provided.

Considering the possible enzymatic activities of WreU, WreQ and WreG, two hypotheses were formed for the initial step of O-antigen biosynthesis in *R. etli* CE3 as described in Chapter One· L. The overall aim of this chapter is to understand how O-antigen biosynthesis is initiated in *R. etli* CE3. Specific aims and reasoning are:

1. To determine whether WreU is the initiating GTase and understand its nucleotide sugar substrate specificity. If WreU can use both UDP-QuiNAc and UDP-ADHexu as substrate but prefers UDP-QuiNAc, hypothesis 1 may be correct; If WreU only uses UDP-ADHexu as substrate, hypothesis 2 is favored.
2. To determine whether WreQ can reduce Und-PP-ADHexu to Und-PP-QuiNAc. If it cannot, then hypothesis 1 is favored in which WreQ is a UDP-ADHexu 4-reductase, and hypothesis 2 may not be correct. If it can, then WreQ is a Und-PP-ADHexu 4-reductase and hypothesis 2 is possible.
3. To determine whether WreG is a mannosyltransferase and understand its specificity to the acceptor substrate. In both hypotheses, WreG may transfer Man to the Und-PP-QuiNAc acceptor, but it should be able also to transfer Man to Und-PP-ADHexu. However, if WreG only transfers Man to the Und-PP-ADHexu acceptor, it will disprove hypothesis 1 and favor hypothesis 2. Specifically, it will indicate that WreQ reduces ADHexu to QuiNAc after the transfer of Man to ADHexu.

Besides *in vitro* methods, the roles of proteins (enzymes) can be investigated *in vivo* through *R. etli* CE166 (*wreQ*::Tn*5* mutant) suppression assays. The low amount of LPS I in CE166 can be brought to near WT level by introducing into this mutant a DNA fragment of the

chromosomal O-antigen genetic cluster on a plasmid. This DNA fragment did not contain the *wreQ* gene and the QuiNAc was still absent in the suppression strain. It was concluded that suppression was due to the multicopy dosage of a gene or genes (35). According to hypothesis 1 and 2, the two enzymes whose activities are most probably affected in CE166 are WreU and WreG. However, the proposed mechanisms for the deficiency of O-antigen biosynthesis and the predicted methods for suppression are different between the two hypotheses, which are discussed below:

Hypothesis 1 proposes that WreU prefers to use UDP-QuiNAc as substrate to initiate O-antigen biosynthesis, although it can also use UDP-ADHexu for the same purpose in the *wreQ*::Tn*5* mutant strain CE166 in which conversion from UDP-ADHexu to UDP-QuiNAc is blocked. Thus, WreU definitely causes the deficiency in O-antigen biosynthesis under this hypothesis. WreG may or may not prefer the acceptor Und-PP-QuiNAc to Und-PP-ADHexu for adding mannose, and it may or may not contribute to the deficiency in O-antigen biosynthesis. Anyhow, under hypothesis 1, CE166 is expected to be suppressed by introducing multiple copies of the *wreU* gene, as this is expected to increase the concentration of Und-PP-ADHexu, which is used by WreG to continue O-antigen biosynthesis.

Hypothesis 2 proposes that WreU preferentially uses UDP-ADHexu as substrate. So, the activity of WreU is not expected to be affected in the *wreQ*::Tn*5* mutant strain CE166. The deficiency in O-antigen biosynthesis should have nothing to do with WreU under this hypothesis, but instead may be caused by WreG as it may prefer the acceptor Und-PP-QuiNAc to Und-PP-ADHexu for adding mannose. In this case, the activity of WreG is definitely affected in CE166 in which ADHexu is not reduced to QuiNAc and it is expected that CE166 can be suppressed by introducing multiple copies of the *wreG* gene.

A variation derived from hypothesis 2 is that WreG prefers to add mannose to Und-PP-ADHexu before the WreQ-catalyzed reduction of ADHexu to QuiNAc. Accordingly, the activity of WreG should not be affected in *wreQ*::Tn*5* mutant strain CE166, and the deficiency in O-

antigen biosynthesis may be caused by a downstream enzyme that recognizes QuiNAc. In this scenario, CE166 is not expected to be suppressed by introducing either the *wreG* or the *wreU* gene.

The *in vivo* suppression of the *wreQ*::Tn*5* mutant strain CE166 with *wreU* or *wreG* genes is expected to provide additional layers of evidence to determine which hypothesis is correct in depicting the initial step of O-antigen biosynthesis in *R. etli* CE3.

B. Investigation of the GTase activity of WreU and its nucleotide sugar substrate specificity

The *wreU* gene was amplified from the chromosomal DNA of *R. etli* CE3 and cloned into pET15b vector. The resultant plasmid encodes a WreU protein with an amino-terminal six-histidine (His_6) tag. The modified *wreU* gene (*His_6-wreU*) was determined to be functional *in vivo* due to its ability to complement an *R. etli wreU*::Km mutant strain CE566 (Fig. 29). The WreU protein is predicted to possess one transmembrane region near its amino terminus. In accordance with this prediction, all His_6-WreU protein expressed in *E. coli* located to the cell membrane fraction. Attempts to purify His_6-WreU from the membrane fraction were not successful. Several other constructs were made in order to facilitate WreU purification, including placing a His_6-tag on the carboxy-terminus of WreU, truncating the predicted amino-terminal transmembrane segment or constructing protein fusions. But none of them yielded WreU protein that can be purified. Thus, the His_6-WreU-containing membrane fraction was used as WreU crude enzyme in the *in vitro* studies.

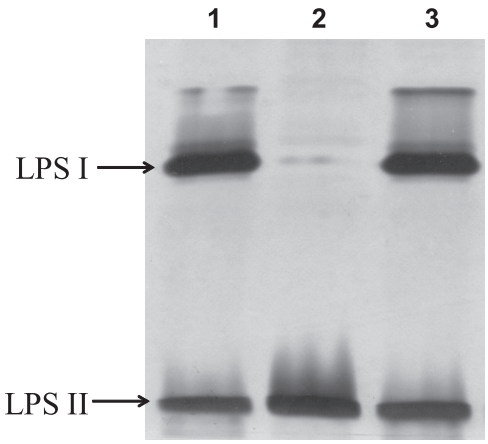


Figure 29. Complementation of the *R. etli wreU*::Km mutant strain CE566 with *His$_6$-wreU*. Silver-stained SDS-PAGE (18% gel) of LPS samples from whole-cell lysates. Lanes: lane 1, CE3 (wildtype); lane 2, CE566 (*wreU*::Km); lane 3, CE566/pLS22 (*wreU*::Km/*His$_6$-wreU*).

The WreU enzyme assay was carried out as a "one-pot" reaction. The putative lipid substrate of WreU, Und-P, was synthesized *in situ* from undecaprenol (Und-OH) and adenosine triphosphate (ATP) by a polyprenol kinase, still known by the misnomer diacylglycerol kinase (DGK), as described by Hartley *et al*. 2008 (62). The nucleotide sugar substrates tested in the reactions were UDP-GlcNAc (commercially available), synthesized UDP-ADHexu (in the WbpM reaction) and UDP-QuiNAc (in the WbpM-WreQ reaction) as described in Chapter Two. After removing protein contents the WbpM reaction product is pure UDP-ADHexu (Fig. 30, lane 2). However, due to incomplete conversion by WreQ, the WbpM-WreQ reaction product is a mixture of UDP-ADHexu and UDP-QuiNAc (Fig. 30, lane 4). Besides using reaction products directly, the UDP-sugars were also purified by TLC (Fig. 30, lane 3 and 5) and used in the WreU assay. All WreU assay reactions were incubated at 30 °C for 1 hour and terminated by adding chloroform-methanol/3:2, followed by extraction into organic and aqueous phases to separate the organic-soluble compounds (such as lipids) and water-soluble compounds (such as nucleotide sugars). The product of the WreU catalyzed reaction is expected to be an undecaprenyl pyrophosphate-linked sugar (Und-PP-sugar). Such a compound is expected to be in the organic phase after extraction.

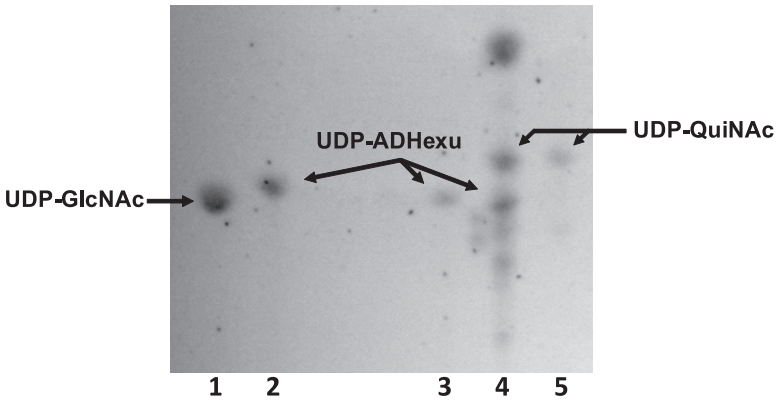


Figure 30. Thymol-stained TLC plate showing the nucleotide sugar substrates for use in the ^{32}P-WreU enzyme assay. Lanes: lane 1, pure UDP-GlcNAc, 10 nmol; lane 2, WbpM reaction product, 5.2 nmol; lane 3, purified UDP-ADHexu, 2.7 nmol; lane 4, WbpM-WreQ reaction product, measurement of UDP-sugar content in this sample is not accurate due to the presence of NAD and NADH; lane 5, purified UDP-QuiNAc, 2.4 nmol.

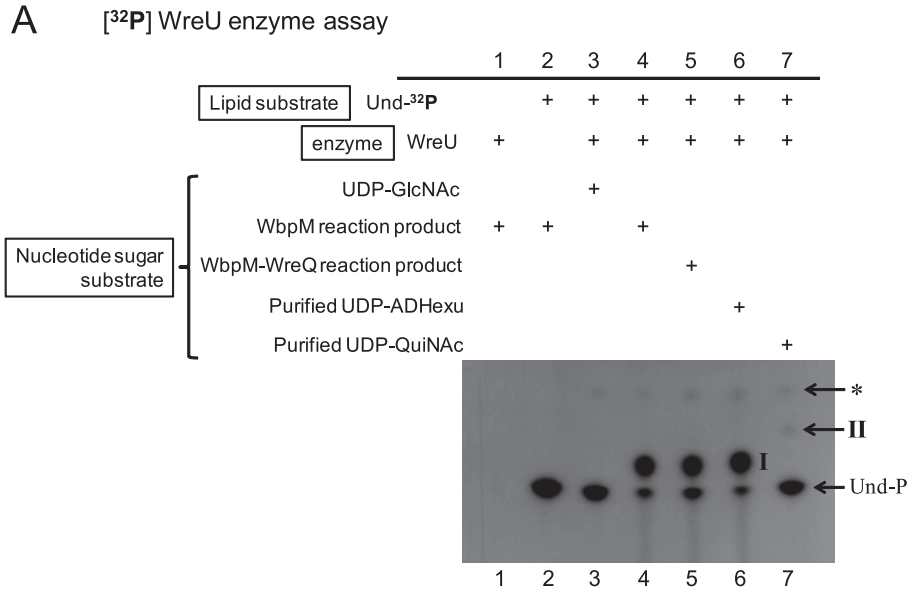
Radioisotope labeling was used in the WreU enzyme assay for detecting compounds of interest by autoradiography after separating them on TLC. Two different radioisotopes were used: phosphorus-32 (^{32}P) and tritium (^{3}H). In the ^{32}P assay, the phosphate group of the lipid substrate Und-P was labeled with ^{32}P by using ATP (γ-^{32}P) in its synthesis. As a result, compounds containing the Und-^{32}P moiety were detected; in the ^{3}H assay, nucleotide sugar substrates were labeled with ^{3}H by using UDP-[^{3}H]GlcNAc in their synthesis. Consequently, only compounds with the ^{3}H-labeled sugar moiety were detected.

The result of ^{32}P-WreU assay is shown in Fig. 31A. A Und-PP-sugar candidate (compound I) was detected in all reactions that contained UDP-ADHexu (Fig. 31A, lane 4-6). No compound with similar migration property was detected in reactions with pure UDP-GlcNAc as nucleotide-sugar substrate (Fig. 31A, lane 3). When purified UDP-QuiNAc was used as nucleotide substrate, a faint spot representing another Und-PP-sugar candidate (compound II) was observed (Fig. 31A, lane 7). Compound I and II may be Und-PP-ADHexu and Und-PP-QuiNAc, respectively. When quantified with phosphorimager, the compound I spots show a 30-fold higher intensity than the compound II spot, suggesting that the WreU catalysis is 30-fold faster with UDP-ADHexu as substrate than UDP-QuiNAc. Importantly, the control reactions confirmed that

the production of Und-PP-sugar compounds required the lipid substrate Und-P (Fig. 31A, lane 1) and the enzyme WreU (Fig. 31A, lane 2).

In the ^{3}H-WreU assay, radioactivity can be detected in the organic phase only if the ^{3}H-sugar moiety is transferred onto the lipid carrier. Thus, this approach provides a more direct observation of the WreU GTase activity. The input ^{3}H-labeled nucleotide-sugar substrates were shown in Fig. 31B (*top* panel). ^{3}H-labeled UDP-QuiNAc was not purified for this assay. The Und-PP-ADHexu candidate (compound I) was observed in a reaction that was supplied with pure UDP-ADHexu from the WbpM reaction (Fig. 31B*, bottom* panel organic phase samples, lane 3). The same compound was also observed in the reaction supplied with the WbpM-WreQ reaction mixture that contained both UDP-ADHexu and UDP-QuiNAc (lane 4). However, the Und-PP-QuiNAc candidate (compound II) was not detected in this reaction, probably because its amount is below detection in our assay. Furthermore, no Und-PP-sugar candidate was observed in the reaction that contained pure UDP-GlcNAc (lane 2) and in the reaction that is devoid of WreU (lane 1). The aqueous phase samples were also analyzed by TLC (Fig. 31B, *bottom* panel aqueous phase samples). Nucleotide sugars that did not participate in the reaction were detected in lane 1 and 2, in accordance with no WreU activity in these reactions. No UDP-ADHexu was detected in lane 3 and 4, consistent with the idea that UDP-ADHexu was used as the substrate by WreU. UDP-QuiNAc was detected in lane 4, suggesting that none or very little of it was used by WreU.

In conclusion, the *in vitro* enzyme assay confirmed the GTase activity of WreU and revealed the nucleotide substrate specificity of WreU: UDP-ADHexu is the preferred substrate, UDP-QuiNAc is a much less favored substrate, and UDP-GlcNAc is not a substrate.


A
[32P] WreU enzyme assay
1 2 3 4 5 6 7
Lipid substrate
Und-32P + + + + + +
enzyme
WreU + + + + + +
Nucleotide sugar substrate
UDP-GlcNAc +
WbpM reaction product + + +
WbpM-WreQ reaction product +
Purified UDP-ADHexu +
Purified UDP-QuiNAc +
*
II
I
Und-P
1 2 3 4 5 6 7

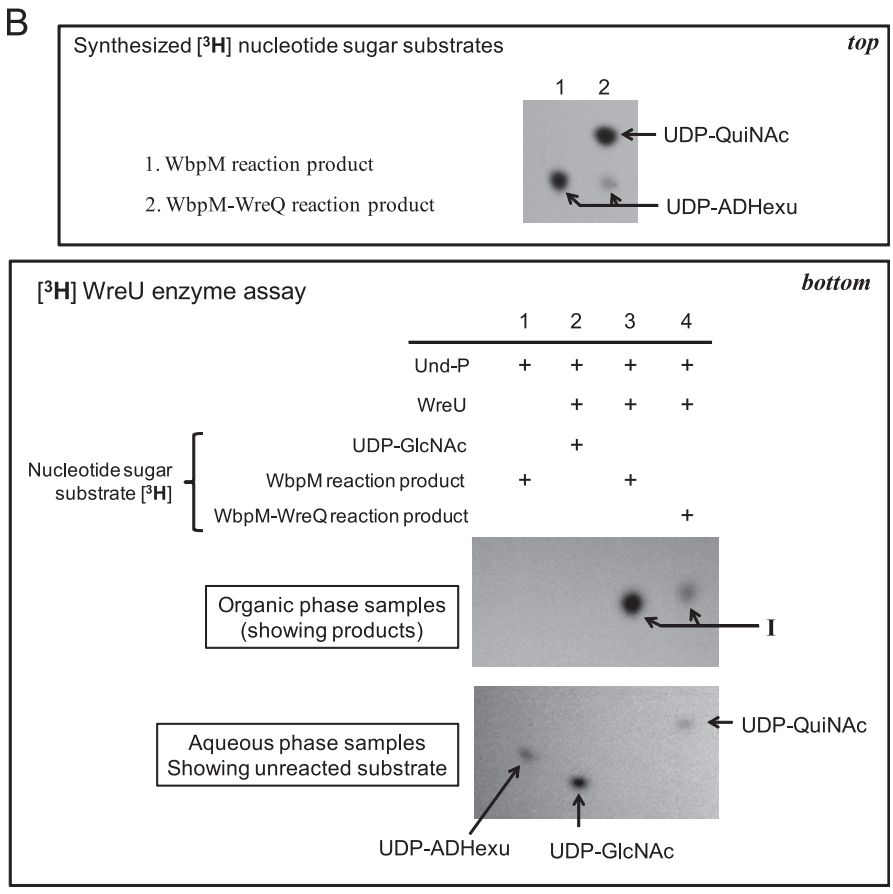

B
Synthesized [3H] nucleotide sugar substrates
top
1 2
1. WbpM reaction product
2. WbpM-WreQ reaction product
UDP-QuiNAc
UDP-ADHexu
[3H] WreU enzyme assay
bottom
1 2 3 4
Und-P + + + +
WreU + + +
Nucleotide sugar substrate [3H]
UDP-GlcNAc +
WbpM reaction product + +
WbpM-WreQ reaction product +
Organic phase samples (showing products)
I
Aqueous phase samples Showing unreacted substrate
UDP-QuiNAc
UDP-ADHexu
UDP-GlcNAc

Figure 31. Analysis of the GTase activity and nucleotide sugar substrate specificity for *R. etli* WreU. (A) Autoradiograms of TLC-separated WreU reactions with radioisotope ^{32}P-labeled lipid substrate Und-P. Lanes: lane 1, no lipid substrate Und-P by omitting the kinase DGK for its synthesis (negative control #1); lane 2, no WreU crude enzyme added to the reaction (negative control #2); lanes 3-7, WreU tested against different nucleotide sugar substrates as shown in the table above the figure. (B) Autoradiogram of TLC-separated WreU reactions with radioisotope ^{3}H-labeled nucleotide sugar substrates. *top* panel, input nucleotide sugar substrates shown by an autoradiogram of TLC-separated UDP-sugar compounds. Identification of each compound has been shown previously by NMR. *bottom* panel, Autoradiograms of TLC-separated compounds in the organic phase and aqueous phase of the WreU reaction. Lanes: lane 1, no WreU (negative control); lanes 2-4, WreU tested against different nucleotide sugar compounds. Compounds: *I*, Und-PP-ADHexu; *II*, Und-PP-QuiNAc; *, a compound derived from Und-P in the presence of *E. coli* membrane, based on its Rf value with a different TLC solvent (not shown), it is possible to be Und-PP (63).

C. WreQ can reduce ADHexu to QuiNAc on the Und-PP linkage

The foregoing result with WreU strongly supported hypothesis 2. According to that hypothesis, the reduction of ADHexu to QuiNAc occurs on the Und-PP linkage. To test whether WreQ can reduce Und-PP-ADHexu to Und-PP-QuiNAc, a WreU-WreQ coupled enzyme assay was carried out.

Like the WreU assay, the WreU-WreQ coupled assay was carried out with both ^{32}P-labeling (Fig. 32A) and ^{3}H-labeling (Fig. 32B). Und-PP-ADHexu (compound I) was produced in the WreU reaction (Fig. 32A, lane 1; Fig. 32B, lane 2), which served as the substrate of WreQ. Since WreQ is a reductase, a reducing agent should be required for its activity and NADH is used in this reaction. The activity of WreQ was tested as follows: When both WreQ and NADH were added to the WreU reaction mixture, compound I was completely converted to a faster-moving compound (compound II) (Fig. 32A, lane 4; Fig. 32B, lane 5). When NADH alone was added without WreQ, no conversion occurred (Fig. 32A, lane 3; Fig. 32B, lane 4); When WreQ was added but NADH was omitted, only a very small amount of compound II was produced (Fig. 32A, lane 2; Fig. 32B, lane 3). Such observations suggested that the conversion from compound I to compound II is a reduction reaction catalyzed by WreQ with NADH. Since WreQ was determined to reduce ADHexu stereospecifically in the D-gluco configuration to QuiNAc in

Chapter Two, compound II is very likely to be Und-PP-QuiNAc. The production of the small amount of compound II by WreQ in the absence of NADH may be explained by: first, there was contaminating NADH present in the reaction either from the crude WreU enzyme (membrane) or the purified WreQ enzyme; second, the amount of WreQ enzyme used in this reaction was in excess and the reaction was allowed to proceed much longer than needed (the evidence is provided in Fig. 33). Importantly, no compound II was produced if WreU was omitted (Fig. 32B, lane 1), which confirmed that compound II must have been converted from the WreU reaction product (compound I).

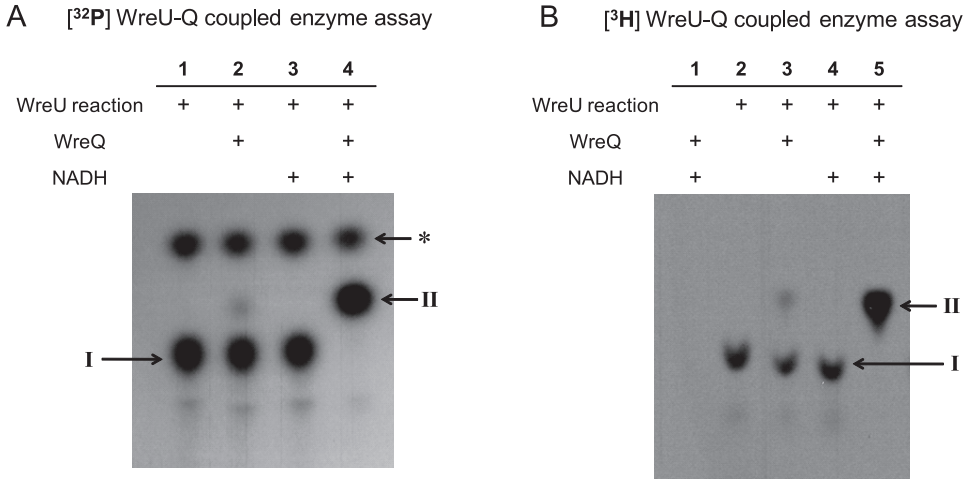


Figure 32. Autoradiograms of TLC-separated compounds in the WreQ enzyme assay organic phase. (A) WreQ enzyme assay with ^{32}P-labelling. Lanes: lane 1, WreU reaction only; lane 2, WreQ added to the WreU reaction; lane 3, NADH added to the WreU reaction; lane 4, both WreQ and NADH added to the WreU reaction. (B). WreQ enzyme assay with ^{3}H-labelling. Lanes: lane 1, WreQ and NADH added to a WreU assay control reaction (no WreU); lanes 2-5, same setup as lanes 1-4 in panel A. Compounds: *I*, Und-PP-ADHexu; *II*, Und-PP-QuiNAc; *, Und-PP (see Fig. 31 legend).

The result of the WreU-WreQ coupled assay suggested that WreQ has Und-PP-ADHexu reductase activity and its product is most likely Und-PP-QuiNAc. In this assay, the amount of WreQ enzyme added and the time allowed for the WreQ reaction were based on previous experiments with the UDP-linked substrate (UDP-ADHexu). To test whether the reduction of the

lipidated substrate (Und-PP-ADHexu) is faster and requires less WreQ enzyme, the WreU-WreQ coupled reactions were performed with serially diluted WreQ and the WreQ reaction was only allowed to proceed for one minute (instead of one hour in the experiment shown in Fig. 32).

The result showed that the amount of Und-PP-ADHexu (compound I) converted to Und-PP-QuiNAc (compound II) in one minute gradually increases with decreasing dilution of WreQ (from 10^{-6} to 10^{-2}) (Fig. 33A, lane 2-6). At 10^{-2}, the conversion was already almost completed (lane 6). Further increase of the amount of WreQ achieved the same result (Fig. 33, lane 7 and 8). The negative control (0 minute) indicated that the method to stop the reaction was effective (lane 1). Kinetic information was estimated from the reactions with 10^{-3} and 10^{-4} diluted WreQ enzyme (Table 2). For comparison, the WreQ-catalyzed UDP-ADHexu reduction reactions were carried out and terminated at different time points (Fig. 33B). The data obtained with one-minute reaction time (Fig. 33B, lane 2) was used for calculation of the reaction rate (Table 2). The result suggested that WreQ catalyzes the reduction of ADHexu to QuiNAc orders of magnitude faster when it is linked to Und-PP rather than UDP.

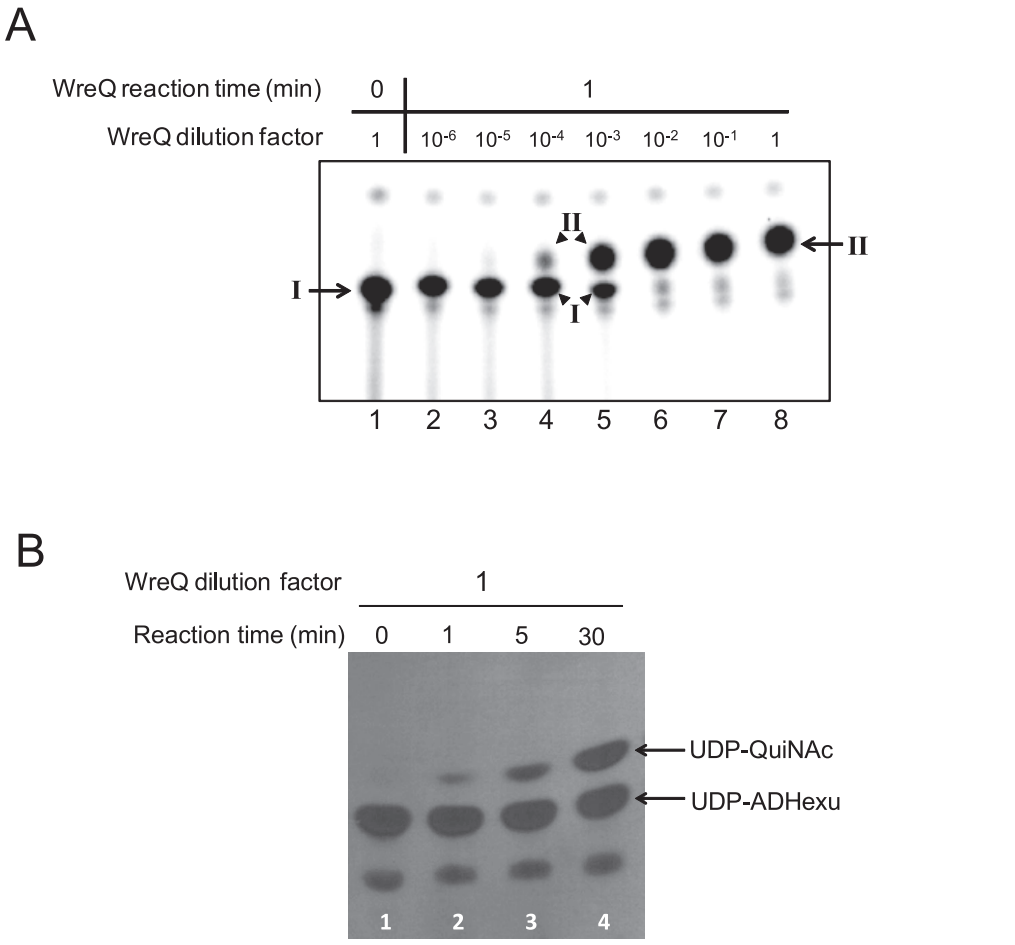


Figure 33. Autoradiograph of TLC-separated WreQ enzyme assay reactions with different substrates. (A) WreQ catalysis with the lipidated substrate. WreQ and NADH were added to the WreU reaction mixture (^{32}P) and allowed to proceed for 1 minute before terminating the reactions by adding chloroform-methnol/3:2 (Solvent I) and votex. Lanes: lane 1 is a 0 minute control in which solvent I was added before WreQ; lane 2-8 are groups that contain serial diluted WreQ, from 10^6 to 1 (undiluted). Compounds: I, Und-PP-ADHexu; II, Und-PP-QuiNAc. (B) WreQ catalysis with UDP-ADHexu. WreQ and NADH were added to the WbpM reaction (^{3}H), in which UDP-ADHexu was produced. The WreQ catalysis was allowed for different amount of time before terminated by boiling. Lanes: lane 1, 0 min; lane 2, 1 min; lane 3, 5 min; lane 4, 30 min.

Table 2. Comparing the estimated enzymatic activities of WreQ with the two substrates, UDP-ADHexu and Und-PP-ADHexu.

Substrate conversion was calculated from the TLC result shown in Fig. 33. The intensities of spots were measured by ImageQuantTL software for ^{32}P spots in Fig. 33A and by ImageJ software for ^{3}H spots in Fig. 33B.

Substrate	Substrate Concentration [S] (μM)	WreQ Concentration [E] (μM)	Substrate Conversion (%)	Product concentration [P] (μM)	Reaction Time t (min)	Reaction Rate V=[P]/t (μM/min)	Turnover Rate$^{\gamma}$ V/[E] (min^{-1})	Turnover/[S]$^{\delta}$ ($\mu M)^{-1}$(min^{-1})
UDP-ADHexu$^{\alpha}$	500	28	13.4	66.9	1	66.9	2.39	0.0048
Und-PP-ADHexu$^{\beta}$	0.33	0.0028	62.5	0.206	1	0.206	74	224
	0.3	0.00028	14	0.042	1	0.042	150	500

$^{\alpha}$ UDP-ADHexu was synthesized *in situ* from WbpM reaction with 0.5 mM UDP-GlcNAc. The conversion to UDP-ADHexu is complete.
$^{\beta}$ Und-PP-ADHexu was synthesized *in situ* from DGK-WreU reaction. Its concentration was calculated from quantified radioactivity counts by phosphorimager.
$^{\gamma}$ Differs from standard turnover number (k_{cat}) by not using (or at least not knowing) saturating [S].
$^{\delta}$ The "turnover rate" in this table underestimates the difference in catalytic efficiencies, whereas the "turnover/[S]" most likely overestimate the difference.

D. WreG adds the second O-antigen sugar mannose to the WreQ reaction product Und-PP-QuiNAc

To obtain the WreG protein for testing its enzymatic assay *in vitro*, the *wreG* gene was amplified from chromosomal DNA of *R. etli* CE3 and cloned into the pET21b vector. The resultant plasmid encodes a WreG protein with a carboxy-terminal His_6 tag. The *wreG-His*$_6$ gene was determined to be functional *in vivo* due to its ability to complement an *R. etli wreG*::Tn*5* mutant strain CE358 (Fig. 34). The WreG-His_6 protein overexpressed in *E. coli* was found to locate in both the soluble fraction and the membrane fraction. An attempt to purify WreG-His_6 from the soluble fraction was not successful, but it can be purified from the membrane fraction after solubilizing membranes with Triton X-100. It is possible that WreG naturally associates with the membrane despite having no predicted transmembrane segment. In the enzyme assay, WreG was supplied either as WreG-His_6-expressing *E. coli* membrane fraction (crude) or as purified protein.

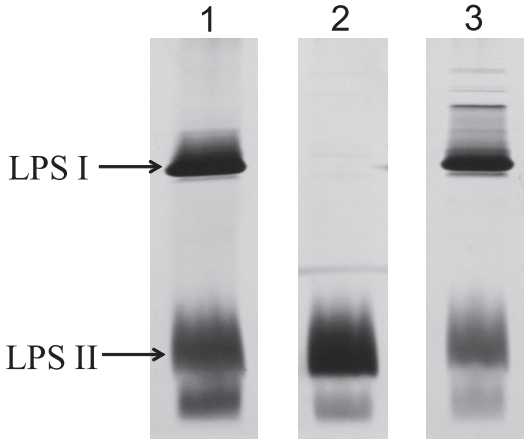


Figure 34. Complementation of the *R. etli wreG*::Tn*5* mutant strain CE358 with *wreG-His*$_6$. Silver-stained SDS-PAGE (18% gel) of LPS samples from whole-cell lysates. Lanes: lane 1, CE3 (wildtype); lane 2, CE358 (*wreG*::Tn*5*); lane 3, CE358/pTL63 (*wreG*::Tn*5*/*wreG-His*$_6$).

To test the acceptor substrate preference of WreG, the WreG enzyme assay was coupled to either the WreU reaction or the WreU-WreQ coupled reaction. The two potential acceptors,

Und-PP-ADHexu (compound I) and Und-PP-QuiNAc (compound II), were produced in the WreU reaction and the WreU-WreQ coupled reaction, respectively. The predicted nucleotide sugar substrate of WreG is GDP-mannose and it was added to all the reactions. Results were analyzed by TLC with two different solvents (Fig. 35). The production of both acceptors was confirmed in parallel control reactions (Fig. 35, lane 2 and 3). When WreG (crude) was incubated with the WreU reaction, Und-PP-ADHexu (compound I) remained almost unchanged (compare lane 4 to lane 2 in Fig. 35). In contrast, when WreG (crude) was incubated with the WreU-WreQ reaction, Und-PP-QuiNAc (compound II) was almost completely converted to a slower-moving compound (compound III) (compare lane 5 to lane 3 in Fig. 35). Compound III is very likely to be Und-PP-QuiNAc-Man, a lipid-linked disaccharide resulting from the GTase activity of WreG. Because compound III migrated at the same place as Und-P in TLC solvent A (Fig. 35, top panel), another solvent (solvent B) was used, which could separate compound III from other radiolabeled compounds (Fig. 35, bottom panel). Purified WreG protein alone also showed GTase activity (Fig. 35, lane 6), albeit much lower compared to the crude enzyme (compare lane 6 to lane 5), as revealed by partial conversion of compound II to compound III.

The result of this assay provided strong evidence that WreG is a mannosyltransferase for adding the second O-antigen sugar mannose in *R. etli* CE3. Furthermore, WreG recognized the acceptor Und-PP-QuiNAc but not Und-PP-ADHexu *in vitro*, suggesting that the reduction of ADHexu to QuiNAc by WreQ should occur first. Combining the results of *in vitro* enzyme assays in this chapter, the order of reactions *in vivo* should be WreU→WreQ→WreG, as proposed in hypothesis 2.

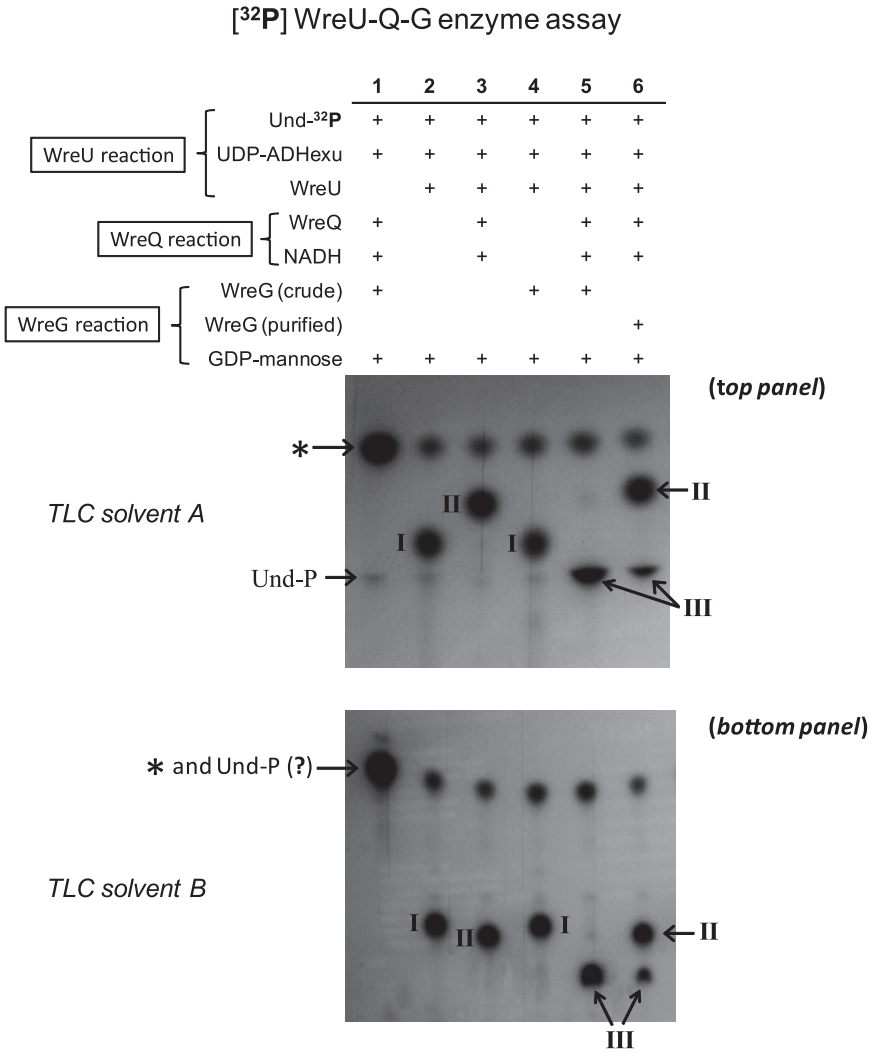


Figure 35. Autoradiograms of TLC-separated WreU-WreQ-WreG coupled enzyme assay organic phase compounds. Dark spots represent ^{32}P-labeled compounds. Samples from six reactions were analyzed by TLC with two different solvents. *top panel*, TLC solvent A: 2-propanol/ammonium hydroxide/water, 6:3:1. *bottom panel*: TLC solvent B: chloroform/methanol/water, 65:25:4. Lanes: lane 1, WreU enzyme was not added; lane 2, only WreU reaction, no WreQ and WreG reaction; lane 3, WreU and WreQ reaction, no WreG reaction; lane 4, WreU reaction plus WreG enzyme, but no WreQ reaction; lane 5. WreU, WreQ and WreG coupled reaction, with crude WreG enzyme (membrane fraction); lane 6, WreU, WreQ and WreG coupled reaction, with purified WreG enzyme. Note that the WreG substrate GDP-mannose was added to each sample. Compounds: *I*, Und-PP-ADHexu; Compound *II*, Und-PP-QuiNAc; Compound *III*, Und-PP-QuiNAc-Man; *, Und-PP (see Fig. 31 legend).

E. *R. etli wreQ*::Tn*5* mutant can be suppressed by introducing multiple copies of *wreG* gene, but not *wreU*.

Observations of *in vitro* reactions described in this chapter to this point are as predicted by hypothesis 2 (see Discussion below). Phenotypes of the *wreQ*::Tn*5* mutant strain (CE166) that fit this hypothesis have been mentioned earlier. As noted in the introduction, additional evidence in favor of this hypothesis may be provided by analyzing the possibility of suppression of the *wreQ* mutant phenotype when multiple copies of other genes are introduced into this mutant strain. The *wreU* and *wreG* genes used for this analysis were *His_6-wreU* and *wreG-His_6*, the same genes encoding protein used in the *in vitro* studies. The two genes were individually cloned into pFAJ1708 vector (51) and the resulting plasmids pLS22 (*His_6-wreU*) and pTL63 (*wreG-His_6*) were independently introduced into strain CE166, causing multicopy dosage of the *wreU* or the *wreG* gene.

Suppression of the *wreQ*::Tn*5* mutant was analyzed by observing the LPS profiles from the suppression strains on SDS-PAGE (Fig. 36). *His_6-wreU* did not suppress the *wreQ*::Tn*5* mutant as the amount of LPS I (O-antigen-containing LPS) in the suppression strain (Fig. 36, lane 3) was unchanged from CE166 (Fig. 36, lane 2), which was much lower than the wildtype (Fig. 36, lane 1). In contrast, *wreG-His_6* suppressed the *wreQ*::Tn*5* mutant by restoring its LPS I amount to near wildtype level (Fig. 36, lane 4). This result indicates that the activity of WreG was affected in the *wreQ*::Tn*5* mutant while the activity of WreU was not affected. It supported hypothesis 2 and is consistent with the conclusion from the *in vitro* studies.

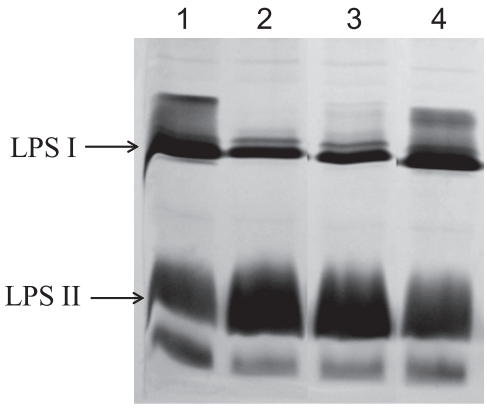


Figure 36. Suppression of *R. etli wreQ*::Tn*5* mutant strain CE166 by introducing multiple copies of *wreU* or *wreG* gene. Lanes: lane 1, CE3 (*R. etli* wildtype strain); lane 2, CE166 (*wreQ*::Tn*5*); lane 3, CE166/pLS22, strain CE166 harboring pLS22 plasmids expressing His_6-WreU; lane 4, CE166/pTL63, strain CE166 harboring pTL63 plasmids expressing WreG-His_6.

F. Discussion

The result of the *in vitro* enzyme assay reported in this chapter confirmed that WreU is the initiating GTase for O-antigen biosynthesis in *R. etli* CE3. Enzymes catalyzing the initiation reaction have been grouped into two distinct families: polyisoprenyl-phosphate *N*-acetylhexosamine-1-phosphate (PNPT) and polyisoprenyl-phosphate hexose-1-phosphate (PHPT) transferases. PNPTs include prokaryotic and eukaryotic members while PHPTs are restricted to bacteria (64). The classification is mainly based on the predicted topology of the proteins and their substrate preference. PNPT family members are represented by WecA, which uses UDP-*N*-acetyl-Hexosamine substrates. PHPT family members are represented by WbaP, which uses UDP-Hexose substrates (64). Based on the predicted topology and the homology to $WbaP_{CT}$, WreU apparently belongs to the PHPT family. However, the nucleotide sugar substrate of WreU determined *in vitro* was UDP-ADHexu (a 4-keto-6-deoxy derivative of UDP-GlcNAc), which belongs to the UDP-*N*-acetyl-Hexosamine category. Thus, WreU represents a subgroup of the PHPT enzyme family. Other characterized enzymes that belong to this subgroup are: WbcP from *Aeromonas hydrophila* AH-3, which is a WbaP homolog and is a UDP-*N*-acetyl-galactosamine (UDP-GalNAc):Und-P GalNAc-1-P transferase (60), and PglC from *Campylobacter jejuni* NCTC

11168, which is homologous to $WbaP_{CT}$ and is a UDP-2,4-diacetamido-2,4,6-trideoxy-α-D-glucopyranose (UDP-Bac):Und-P Bac-1-P transferase (65).

The lipid substrate of WreU in the *in vitro* reactions is an undecaprenol lipid carrier (Und-P). However, the exact form of bactoprenol in *R. etli* CE3 is not known. A dodecaprenol lipid carrier (C_{60}) has been reported to be involved in a step of core synthesis in *Rhizobium leguminosarum* 3841 and it was also present in *Sinorhizhobium meliloti* 1021 (66). It is possible that the dodecaprenol lipid carrier also exists in *R. etli* CE3 and is the carrier for O-antigen synthesis.

Three nucleotide sugars were tested as potential substrates for WreU. They are UDP-GlcNAc, UDP-ADHexu and UDP-QuiNAc. In the previous chapter, UDP-QuiNAc was synthesized *in vitro* from UDP-GlcNAc in a two-step reaction via an intermediate UDP-ADHexu, and this was the previously proposed pathway for UDP-QuiNAc biosynthesis *in vivo* (19). In this chapter, WreU was shown to prefer UDP-ADHexu as substrate. This result led to the conclusion that WreU is a UDP-ADHexu:Und-P ADHexu-1-P transferase and strongly suggested that the reduction of ADHexu to QuiNAc occurs on the Und-PP linkage *in vivo* instead of on the UDP linkage as previously proposed.

The reduction of ADHexu to QuiNAc requires 4-reductase activity. In the previous chapter, WreQ exhibited 4-reductase activity in reducing UDP-ADHexu to UDP-QuiNAc *in vitro*. However, this may be an artificial reaction which only occurs *in vitro* with high concentrations of purified enzymes and substrates. In *R. etli*, it is possible that the WreU activity is tightly coupled to the 4,6-dehydratase so that all synthesized UDP-ADHexu was used by WreU for the initiating reaction. In this case, UDP-QuiNAc may never be synthesized in the cell because WreQ does not encounter UDP-ADHexu. Instead, WreQ is expected to reduce ADHexu on the Und-PP linkage. And indeed, WreQ was shown in this chapter to catalyze the reduction of Und-PP-ADHexu to Und-PP-QuiNAc.

In conclusion, biosynthesis of QuiNAc in *R. etli* CE3 is coupled to O-antigen initiation, as depicted in Fig. 37. In the first phase of QuiNAc biosynthesis (4,6-dehydration), the precursor GlcNAc is converted to ADHexu by WreV on the UDP linkage. In the second phase, the product UDP-ADHexu is used by WreU to initiate O-antigen biosynthesis by transferring ADHexu-1-P to Und-P. In the third (final) phase of QuiNAc biosynthesis (4-reduction), ADHexu is reduced to QuiNAc by WreQ on the Und-PP linkage. The product Und-PP-QuiNAc is considered as the first O-antigen intermediate and it serves as acceptor for the next O-antigen sugar.

WreG was confirmed to be a GTase that adds the second O-antigen sugar mannose. The *in vitro* reactions showed that WreG transfers mannose from GDP-mannose to the Und-PP-QuiNAc acceptor, but its activity was slight or undetectable with the Und-PP-ADHexu acceptor. On the one hand, it supported that the reduction of ADHexu to QuiNAc occurs before the addition of the next sugar. On the other hand, mannose should be able to be attached to ADHexu as this occurs in the *R. etli wreQ*::Tn*5* mutant in which ADHexu is not reduced to QuiNAc yet full-length O antigen is synthesized. This activity seems to be conferred by WreG and not another GTase because multiple dosage of *wreG* gene could suppress the deficiency in O-antigen biosynthesis in the *wreQ*::Tn*5* muant. It is possible that WreG could only add mannose to ADHexu *in vivo* and such a reaction requires certain conditions that could not be met *in vitro*. The WreG result *in vitro* gives confidence that the reactions studied *in vitro* in this chapter are true O-antigen synthesis reactions *in vivo*. The conclusion that WreG is the GTase, that mannose is the sugar added, and that QuiNAc is the acceptor fits the O-antigen structure (4) and the genetics presented in previous studies (12).

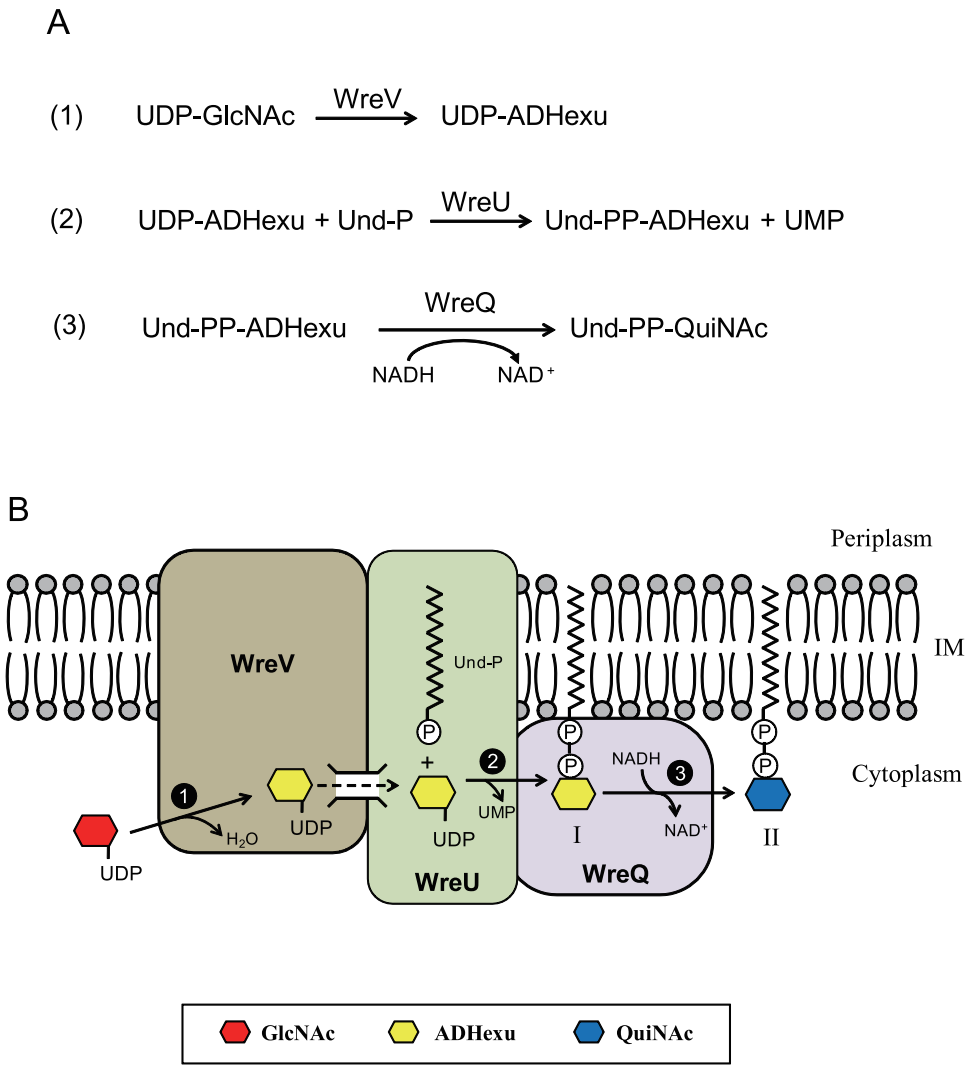


Figure 37. Biosynthesis of QuiNAc is coupled to O-antigen initiation in *R. etli* CE3. (A) Reaction catalyzed by each enzyme in the QuiNAc biosynthesis pathway in *R. etli* CE3. (B) Model of QuiNAc synthesis coordinated with O-antigen initiation. The three phases of QuiNAc synthesis are indicated with numbers: first, UDP-GlcNAc is converted to an intermediate (UDP-ADHexu) by the dehydratase WreV (phase #1); second, ADHexu-1-P is transferred by WreU to the Und-P lipid carrier (phase #2); third and last, the ADHexu moiety is reduced to QuiNAc by WreQ (phase #3). A potential substrate channel for UDP-ADHexu between WreV and WreU is indicated. Compounds: I, Und-PP-ADHexu; II, Und-PP-QuiNAc.

The results of this chapter support the second hypothesis of the introduction. It differs from the first hypothesis mainly in that QuiNAc is synthesized on the Und-PP linkage rather than the UDP linkage. The newly identified pathway of QuiNAc biosynthesis also fits the predicted localization of enzymes involved in this process. The 4,6-dehydratase WreV that converts UDP-GlcNAc to UDP-ADHexu is predicted to be an integral membrane protein and so is the initiating

GTase WreU that transfers ADHexu-1-P to Und-P. The two membrane-located enzymes may have substrate channeling which allows for a more efficient reaction (as shown in Fig. 37B). The WreQ protein may also be membrane associated and it may form a complex with WreV and WreU. This complex would allow efficient synthesis of the first O-antigen intermediate Und-PP-QuiNAc in a confined compartment. Besides being more efficient, such an organization of enzymes may be protective as well. Several 4-keto-6-deoxy sugars have been reported to be unstable (52,57-59). UDP-ADHexu was stable enough *in vitro* to withstand TLC purification and NMR data acquisition, but it is unclear how stable will it be in the cytosol.

CHAPTER FOUR: GENOMIC INVESTIGATION OF QUINAC BIOSYNTHESIS

As demonstrated in the previous chapter, biosynthesis of QuiNAc follows an unconventional pathway in *R. etli* CE3 (Fig. 37). The three enzymes involved in this process, a 4,6-dehydratase, an initiating GTase and a 4-reductase, are encoded by genes *wreV*, *wreU* and *wreQ*, respectively in *R. etli* CE3. There are no biochemical investigations on QuiNAc synthesis in other bacteria. To gain knowledge about the prevalence of the QuiNAc synthesis pathway identified in *R. etli* CE3, a genomic approach was undertaken. Specifically, the homologs of *R. etli* WreV, WreU and WreQ were searched for in other bacterial strains using BLAST.

A. WreU, WreV, WreQ are conserved in *R. etli and R. leguminosarum* strains

In *R. etli* CE3, it is curious that *wreV*, *wreU* and *wreQ* are all located outside of the 28-kb chromosomal O-antigen genetic cluster, where most of the genes required for O-antigen biosynthesis are located. Even these three genes are not all located together: *wreU* and *wreV* are found on plasmid B with *wreV* downstream of *wreU* (20); *wreQ* seems to be randomly situated in the chromosome, bracketed by two cell wall synthesis genes (Fig. 6 (of Chapter One), Fig. 38A) (15,19).

To investigate whether the presence of QuiNAc and its synthesis pathway may be conserved in other strains of *R. etli* and in strains of the closely related species *R. leguminosarum*, genomic sequence of these strains were examined for genes encoding homologs of WreV, WreU and WreQ, and their locations in the genome. Besides *R. etli* CE3, complete genomic sequence is available for *R. etli* bv. *mimosae* str. Mim1, *R. etli* CIAT 652, *R. leguminosarum* bv. *viciae* 3841, *R. leguminosarum* bv. *trifolii* WSM2304, *R. leguminosarum* bv. *trifolii* WSM1325, *R. leguminosarum* bv. *trifolii* WSM1689, and *R. leguminosarum* bv. *trifolii* CB782.

R. etli CE3 *wreU*, *wreV* and *wreQ* gene homologs were found in all these strains with very high sequence conservation (as summarized in Table 3). Moreover, their locations in the

genome of these strains are conserved (Fig. 38). In all cases, the 4-reductase gene *wreQ* is located on the chromosome, while the initial GTase gene *wreU* and the 4,6-dehydratase gene *wreV* are found on a plasmid.

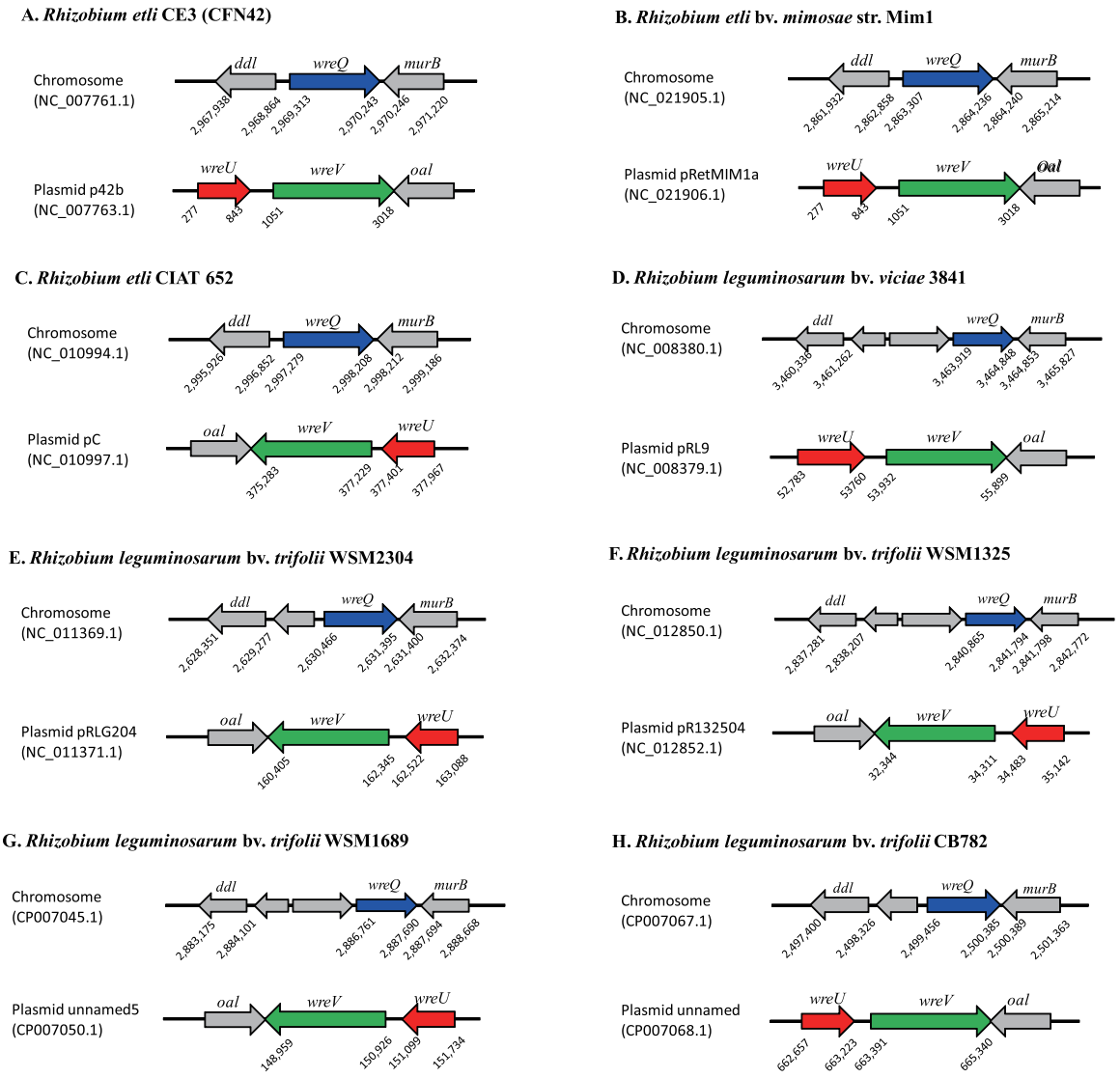


Figure 38. Relative locations of *wreQ*, *wreU* and *wreV* genes and their homologs on the genomes of *Rhizobium* strains. (A-H) Eight *R. etli* or *R. leguminosarum* strains that have complete genomic sequences. *wreQ, wreU* and *wreV* gene homolog are in blue, red, and green, respectively. Other genes are shown in grey: *oal*, O-antigen ligase gene, previously designated *orfL* in *R. etli* CE3 (21); *ddl*, d-alanine--d-alanine ligase gene; *murB*, UDP-*N*-acetylenolpyruvoylglucosamine reductase gene.

Table 3. All *Rhizobium etli* and *Rhizobium leguminosarum* strains that have a complete genomic sequence encode *R. etli* CE3 WreQ, WreV and WreU homolog. Accession numbers are from genome sequencing projects. E-values are from BLAST alignment with genes of *R. etli* CE3.

Bacterial Strain	WreQ homolog		WreV homolog		WreU homolog	
	Accession No.	E-value	Accession No.	E-value	Accession No.	E-value
R. etli CE3	YP_470339	0	YP_471773	0	YP_471772	2e-131
R. etli bv. *mimosae* str. Mim1	YP_008365580	0	YP_008366884	0	YP_008366882	2e-139
R. etli CIAT 652	YP_001979122	0	YP_001984979	0	YP_001984980	1e-130
R. leguminosarum bv. *viciae* 3841	YP_768883	0	YP_765345	0	YP_765344	1e-121
R. leguminosarum bv. *trifolii* WSM2304	YP_002282088	0	YP_002284248	0	YP_002284249	4e-123
R. leguminosarum bv. *trifolii* WSM1325	YP_002976650	0	YP_002978517	0	YP_002978518	3e-124
R. leguminosarum bv. *trifolii* WSM1689	AHF84840	0	AHF88737	0	AHF88738	3e-124
R. leguminosarum bv. *trifolii* CB782	AHG46133	0	AHG48652	0	AHG48651	1e-129

The *wreQ* gene is inserted in a cluster of peptidoglycan biosynthesis genes, between *ddl* (D-alanine-D-alanine ligase gene) and *murB* (UDP-*N*-acetylenolpyruvoylglucosamine reductase gene), and *wreQ* is transcribed in the opposite direction. In *R. leguminosarum* strains, one or two other genes are also inserted in this locus (Fig. 38D-H). They are not functionally related to QuiNAc or O-antigen synthesis.

The *wreU* and *wreV* genes are transcribed in the same direction with *wreU* upstream of *wreV*. The two genes are generally separated by around 200 bp with specific variations among strains. It is unclear whether their transcription is driven by the same promoter or *wreV* has its own promoter. A gene encoding a putative O-antigen ligase is located immediately downstream of *wreV* in all sequenced *Rhizobium* strains, and it is transcribed in an opposite direction. The mutagenesis of this gene in *R. etli* CE3 will be discussed in Chapter Six.

In conclusion, the presence of *wreU*, *wreV*, *wreQ* homologous genes and their conserved location suggest that all of these *Rhizobium* strains may produce QuiNAc as the primer sugar of a polysaccharide built on the lipid carrier Und-P. The polysaccharide is mostly likely O antigen as in *R. etli* CE3 and the conserved genetic cluster containing *wreU*, *wreV* and the putative O-antigen ligase gene seems to support that notion.

Among *R. etli* and *R. leguminosarum* strains that have a completely sequenced genome except *R. etli* CE3, only the structure of the primary O antigen of *R. leguminosarum* 3841 has been solved (67). QuiNAc was found to be a component of that O-antigen. However, the authors concluded in their structural study that the QuiNAc is in the L-configuration, instead of the D-configuration as in *R. etli* CE3 O-antigen. This conclusion is inconsistent with the presence of highly conserved gene homologs in *R. leguminosarum* 3841 (Table 3 and Fig. 38*D*), which predict it to make D-QuiNAc. Moreover, the L-QuiNAc biosynthetic pathway has been demonstrated by *in vitro* enzymatic studies of gene products of *Vibrio cholerae* serotype O37 (68). No gene homologs involved in L-QuiNAc biosynthesis were identified in the genome of *R. leguminosarum* 3841 in a BLAST search. This inconsistency between structural and genomic

data may be explained in several ways: first, *R. leguminosarum* 3841 may use a yet unknown pathway for synthesizing L-QuiNAc; second, the QuiNAc may be actually in D-configuration; and third, D-QuiNAc may be a component of a different O antigen that was not isolated in that structural study.

The O-antigen structure of another *R. leguminosarum* strain, RBL5523, has also been solved (69). It is composed of →4)-α-D-Glc-(1→3)-α-D-QuiNAc-(1→ repeating units. Since genomic sequence is not available for this strain, the presence of *wreU*, *wreV* and *wreQ* homologs could not be verified. However, the presence of QuiNAc in the D-configuration in the O antigen of this strain supports our conclusion that the production of QuiNAc may be conserved in *R. etli* and *R. leguminosarum* strains.

B. The presence of WreV and WreQ homolog in QuiNAc-containing bacteria

QuiNAc has been identified in the O antigen or the capsular polysaccharide of bacterial strains that belong to other species (Table 4). Only three strains among those have a completely sequenced genome, they are *Vibrio vulnificus* M06-24, *Legionella pneumophila* subsp. *pneumophila* str. Philadelphia 1 and *Francisella tularensis* subsp. *tularensis* SCHU S4. Another QuiNAc-containing strain, *Pseudomonas aeruginosa* O6, while not having a complete genomic sequence, has been studied genetically for its B-band O-antigen synthesis and the sequence of the O-antigen genetic cluster is available (43).

Without considering the transfer step, the synthesis of QuiNAc from the precursor GlcNAc requires the two sugar-modifying enzymes, the 4,6-dehydratase (*R. etli* CE3 WreV homolog) and the 4-reductase (*R. etli* CE3 WreQ homolog). The presence of the two enzymes would be predicted in all QuiNAc-containing bacteria unless they synthesize QuiNAc in a completely different way. Indeed, genes encoding the WreV and WreQ homolog were found in all of the four bacterial strains (Table 5).

Table 4. Bacterial strains that have been shown to contain D-QuiNAc in polysaccharide structures.

Species	Strain (Reference)
Aeromonas salmonicida	80204-1 (70)
Burkholderia caryophylli	NCPP 2151 (71)
Francisella tularensis	Strain 15 (72), subsp. *tularensis* SCHU S4, subsp. *holarctica* LVS (73)
Legionella pneumophila	subsp. *pneumophila* str. Philadelphia 1 (74)
Photorhabdus asymbiotica	subsp. *asymbiotica* US-86 and US-87 and subsp. *australis* AU36, AU46, and AU92 (75)
Pseudoalteromonas distincta	KMM 638 (76)
Pseudomonas aeruginosa	O1, O4, O6, O9, O10, O12, O13, O14, O19 (45,77)
Pseudomonas tolaasii	NCPPB2192 (78)
Rhizobium etli	CE3 (4)
Rhizobium leguminosarum	RBL5523 (69)
Vibrio cholerae	H11 (79), O139 (80), O43 (81)
Vibrio vulnificus	M06-24 (82), 6353 (83)

The presence of these two genes may be a conserved feature in all bacteria containing QuiNAc. Conversely, the presence of QuiNAc in bacterial strains could be predicted based on whether they contain the WreV and WreQ homologs. For instance, we predict that *Chlorobium phaeobacteroides* DSM 266, *Herbaspirillum seropedicae* SmR1, *Polaromonas naphthalenivorans* CJ2, as well as several *Rhizobium* and *Pseudomonas* strains not listed in Table 5, produce QuiNAc that would be found if the compositions of their polysaccharides were to be analyzed. More examples are given in Table 6. In most of these bacterial strains, genes encoding the WreV and WreQ homolog are close to each other, indicating that they are functionally related.

Table 5. The presence of WreQ, WreV and initiating GTase homolog in QuiNAc-containing bacterial strains that have a sequenced genome.

Bacterial Strain	WreQ homolog		WreV homolog		Initiating GTase			
					WreU homolog		WbpL homolog	
	Accession No.[b]	E-value	Accession No.	E-value	Accession No.	E-value	Accession No.	E-value
Rhizobium etli CE3 (4)[a]	YP_470339	0	YP_471773	0	YP_471772	2e-131		
Pseudomonas aeruginosa O6 (84)	AAF23991	2e-98	AAF23989	0			AAF23990	0
Vibrio vulnificus M06-24 (82)	YP_004189989	3e-104	YP_004189989	0	YP_004189990	2e-38		
Legionella pneumophila subsp. *pneumophila* str. Philadelphia 1 (74)	YP_094797	1e-77	YP_095000	8e-122			YP_094798	1e-31
Francisella tularensis subsp. *tularensis* SCHU S4 (73)	YP_170399	9e-32	YP_170401	2e-116	YP_170400	1e-60		

[a] References following the name of bacterial strains are studies that identified QuiNAc in that strain.
[b] Accession numbers are from genome sequencing projects, except *Pseudomonas aeruginosa* O6, for which the accessions were associated with Belanger, *et al*., 1999 (43).

Table 6. Bacterial Strains predicted to contain QuiNAc based on the presence of genes encoding homologs of *R. etli* CE3 WreV and WreQ in their genome.

WreQ was used as the query sequence in a BLAST search against the non-redundant protein sequences (nr) database using the blastp algorithm. Bacterial strains selected all meet the following criteria: 1. encode a WreQ homolog with lower e-value than that of WbpV, 7e-91 (the WreQ homolog of *P. aeruginosa* O6); 2. have a fully-sequenced genome; 3. encode a WreV homolog (40 out of 41 strains contain both WreV and WreQ homolog).

Genus	Species	Strain
Chlorobium	*phaeobacteroides*	DSM 266
Colwellia	*psychrerythraea*	34H
Cycloclasticus	*sp.*	P1
Gallionella	*capsiferriformans*	ES-2
Geobacter	*bemidjiensis*	Bem
	sp.	M21
Herbaspirillum	*seropedicae*	SmR1
Herminiimonas	*arsenicoxydans*	
Marinobacter	*aquaeolei*	VT8
Methylobacillus	*flagellatus*	KT
Methylomicrobium	*alcaliphilum*	20Z
Methylotenera	*mobilis*	JLW8
	versatilis	301
Nitrosomonas	*eutropha*	C91
Pelodictyon	*phaeoclathratiforme*	BU-1
Polaromonas	*naphthalenivorans*	CJ2
	sp.	JS666
Pseudomonas	*entomophila*	L48
	fluorescens	Pf0-1
	mendocina	ymp
	monteilii	SB3078
	protegens	CHA0, Pf-5
	putida	BIRD-1, GB-1, HB3267, W619
	stutzeri	DSM 10701, RCH2
Rhizobium	*etli*	CIAT 652, bv. *mimosae* str. Mim1
	leguminosarum	bv. *trifolii* WSM1325, bv. *trifolii* WSM1689, bv. *trifolii* WSM2304, bv. *viciae* 3841
Sideroxydans	*lithotrophicus*	ES-1
Thioalkalivibrio	*sp.*	K90mix
Vibrio	*cholerae*	O1 biovar El Tor str. N16961
Variovorax	*paradoxus*	EPS

C. The presence of an initiating GTase genes next to the *wreQ* and (or) *wreV* homolog

Due to the involvement of an initiating GTase, the QuiNAc synthesis pathway identified in *R. etli* CE3 may only apply to the cases in which QuiNAc is the first sugar transferred to the lipid carrier to initiate polysaccharide synthesis. The definitive evidence for a sugar being the initiating sugar can only be provided by *in vitro* biochemical studies, like in this dissertation. However, QuiNAc has been proposed to be the initiating sugar of the O-antigen repeating unit in *P. aeruginosa* O6 (43). It has not been specifically stated whether QuiNAc is the initiating sugar in other QuiNAc-containing polysaccharides. However, there are no data conflicting with the idea that QuiNAc is always the initiating sugar, either.

In *R. etli* CE3, the O-antigen initiating GTase encoded by *wreU* is involved in QuiNAc synthesis. It transfers a 4-keto intermediate to the lipid carrier Und-P and the intermediate is reduced to QuiNAc on the lipid linkage. In *P. aeruginosa* O6, the *wbpL* gene encodes an initiating GTase that has been proposed to initiate synthesis of the repeating unit of B-band O-antigen. QuiNAc was proposed to be the first sugar of the repeating unit (43) and WbpL was proposed to transfer QuiNAc-1-P from UDP-QuiNAc to the lipid carrier. However, this conclusion is based on genetic studies and WbpL has not been characterized biochemically. WbpL belongs to the PNPT family initiating GTase and its predicted topology is very different from WreU, which belongs to the PHPT family (Fig. 39). In the genome of *P. aeruginosa* O6, *wbpL* is located between the *wbpV* (*wreQ* homolog) and *wbpM* (*wreV* homolog) (Fig. 40B). To investigate whether an initiating GTase gene is always found next to the *wreV* and *wreQ* homolog, WreU of *R. etli* CE3 or WbpL of *P. aeruginosa* O6 was used as query in BLAST searches in QuiNAc-containing bacteria listed in Table 5.

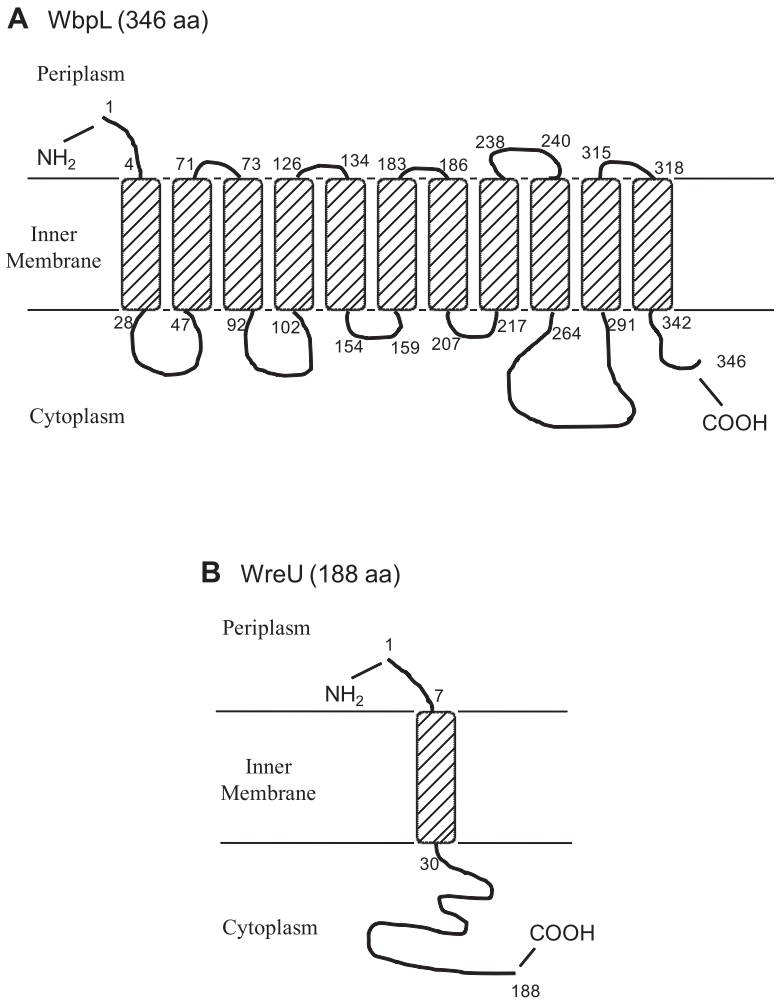


Figure 39. Comparison of predicted topology between initiating GTase WbpL and WreU. The shaded rectangular boxes represent transmembrane (TM) segments. The numbers indicate the amino acid positions of the boundary of each TM domain. (A) *Pseudomonas aeruginosa* O6 WbpL (Accession No.: AAF23990); (B) *Rhizobium etli* CE3 WreU (Accession No.: YP_471772).

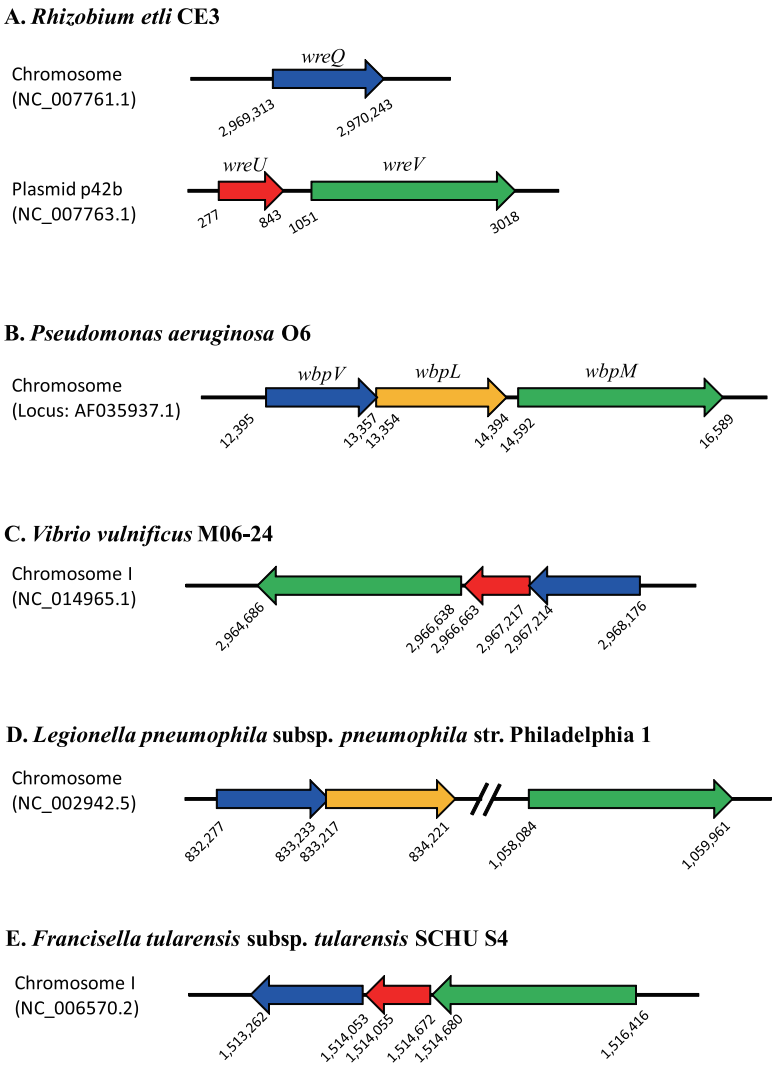


Figure 40. Relative locations of homologous genes in *R. etli* CE3 and the four known genomes of QuiNAc-containing bacterial strains. *wreV*, *wreQ*, *wreU*, *wbpL* and their homologs are shown in green, blue, red, orange, respectively.

The result of the search is also listed in Table 5. *V. vulnificus* M06-24 and *F. tularensis* subsp. *tularensis* SCHU S4 each contains a WreU homolog. *L. pneumophila* subsp. *pneumophila* str. Philadelphia 1 contains a WbpL homolog. The location of the initiating GTase genes were examined together with the *wreV* and *wreQ* homologs in these bacterial strains (Fig. 40). Interestingly, the three genes tend to be contiguous with the initial GTase gene in the middle (Fig. 40B,C,E), or, in the case of *R. etli* CE3 (Fig. 40A) and *L. pneumophila* subsp. *pneumophila* str. Philadelphia 1 (Fig. 40D), two genes are contiguous while the other one is located elsewhere. In

strains predicted to produce QuiNAc (Table 6), it is also common for the three genes or at least two genes to be contiguous (data not shown). This pattern of gene location suggests that these genes may have related functions, that is, they are involved in QuiNAc synthesis and its transfer as the first sugar of a polysaccharide structure.

D. Summary of the findings and speculation

The ubiquity of a sugar ring-modifying 4,6-dehydratase (*R. etli* WreV homolog) and 4-reductase (WreQ homolog) in QuiNAc-containing bacteria suggests that QuiNAc conversion from GlcNAc is conserved. However, there are at least two classes of initiating GTases: WreU or WbpL. The close linkage between an initiating GTase with the two sugar ring-modifying enzymes further suggests that QuiNAc may always be the initiating sugar of a polysaccharide built on the lipid carrier Und-P. It is highly possible that QuiNAc is synthesized in the same way in all QuiNAc-containing bacteria that possess WreU homologs. Although WbpL and its homologs have not been characterized biochemically, they might also catalyze the same reaction as WreU. In our model (Fig. 37B), we speculated that WreU forms a membrane complex with WreV. If WbpL catalyzes the same reaction as WreU, it may also interact with the WreV homolog (WbpM). Given the differences in their topologies, it is hard to be believe that they interact with the 4,6-dehydratase (WreV or WbpM) in the same manner. And this may potentially argue against the necessity of complex formation.

CHAPTER FIVE: DISCUSSION

A. UDP-QuiNAc may be an artificial compound

In the present study, the biosynthesis pathway of QuiNAc was determined. Unlike the previously proposed canonical pathway in which synthesis occurred entirely on the UDP linkage, QuiNAc synthesis actually begins on the UDP-linkage but finishes on the Und-PP linkage (as detailed in Fig. 37). According to this pathway, UDP-QuiNAc may never be synthesized in the cell. The reason for synthesizing QuiNAc in such a manner is unclear. One explanation is that UDP-QuiNAc may be toxic to the cell. Due to structural similarity to UDP-GlcNAc, UDP-QuiNAc may affect many enzymes that are involved in UDP-GlcNAc metabolism and catalysis. It may inhibit the activity of these enzymes, or it may be used as substrate instead of UDP-GlcNAc and affect a downstream event. Bacterial peptidoglycan (PGN) consists of alternating residues of GlcNAc and *N*-acetylmuramic acid (MurNAc)-pentapeptide. UDP-GlcNAc is crucial for PGN biosynthesis as the donor of GlcNAc and the biosynthetic precursor of UDP-MurNAc (85). Thus, cell wall synthesis may be one of the cellular processes affected by UDP-QuiNAc. The QuiNAc biosynthesis pathway determined in this study revealed an approach by which the cell can produce QuiNAc without the synthesis of UDP-QuiNAc.

B. Potential applications of UDP-QuiNAc

UDP-QuiNAc can be produced enzymatically in an *in vitro* reaction as described in Chapter Two. Although it may be an unnatural compound, its availability has many potential applications. It is a useful nucleotide sugar substrate candidate for testing the specificity of glycosyltransferases. In this dissertation, UDP-QuiNAc has been successfully applied in testing specificity of the *R. etli* O-antigen initiating GTase WreU. The finding that UDP-ADHexu was a

preferred substrate over UDP-QuiNAc was crucial in determining the pathway of QuiNAc biosynthesis.

UDP-QuiNAc could also be used in testing the specificity of other initiating GTases. For example, WecA of *E. coli* has been determined *in vitro* to use UDP-GlcNAc as substrate and transfer GlcNAc-1-P to Und-P (26). UDP-QuiNAc is structurally homologous to UDP-GlcNAc, the only difference is at the sixth carbon of the hexose ring ($-CH_3$ for QuiNAc, $-CH_2OH$ for GlcNAc). So, UDP-QuiNAc may be a substrate of WecA or it may be an inhibitor of WecA activity. The specific role of UDP-QuiNAc will be determined by the active site architecture of WecA. The crystal structure of WecA has not been solved, but if this can be achieved, UDP-QuiNAc may help in understanding the nature of substrate specificity.

Another GTase that can be tested with UDP-QuiNAc is MurG, which is involved in PGN synthesis. MurG transfers GlcNAc from UDP-GlcNAc to Und-PP-MurNAc-pentapeptide, which is the second step in synthesizing the Und-PP linked MurNAc-(pentapeptide)-GlcNAc alternating unit of PGN (13). UDP-QuiNAc might affect PGN synthesis through MurG in two ways. First, it may inhibit the activity of MurG thus blocking PGN synthesis. Second, UDP-QuiNAc may be used by MurG as a substrate and added instead of GlcNAc. However, the MurNAc-(pentapeptide)-QuiNAc units may not be efficiently polymerized, which will also affect PGN synthesis. Given the essential role of PGN and the potential deleterious effects of UDP-QuiNAc on its synthesis, it makes sense for QuiNAc to be produced in the anchored, lipid-linked form instead of the soluble, UDP-linked form.

UDP-QuiNAc may also be used for *in vitro* assembly of polysaccharides. If the initial GTase that normally uses UDP-GlcNAc can also use UDP-QuiNAc as substrate, the lipid-linked QuiNAc can be used as acceptor for the following steps of polysaccharide synthesis. By comparison with lipid-linked GlcNAc as acceptor, the importance of the primer (or first) sugar can be deduced. In protein N-glycosylation in eukaryotes, the first sugar of the glycan chain is GlcNAc (86). The initiating GTase (ALG7) that transfers GlcNAc-1-P onto the lipid carrier

dolichol phosphate is homologous to WecA. If this enzyme could allow substitution of GlcNAc with QuiNAc and the rest of glycan chain synthesis is not affected, it provides a way to label the glycan chain with radioisotopes. Specifically, radiolabeled (^{3}H or ^{14}C) UDP-QuiNAc can be used in the assembly of the glycan. Because UDP-QuiNAc is less likely to be used by other enzymes in the cell than UDP-GlcNAc, it is expected to provide a cleaner background.

C. The uniqueness of the QuiNAc synthesis pathway

To date, more than 100 sugar moieties are found in bacterial polysaccharides. Except for a few sugars such as glucose (Glc), galactose (Gal) and *N*-acetylglucosamine (GlcNAc) which are common components of many other structural glycans and are utilized in other housekeeping metabolic functions, most other sugars are synthesized as sugar nucleotides from nucleotide-activated common sugars. Their biosynthetic pathways may comprise one to multiple steps (36). To our knowledge, in all biochemically characterized sugar synthesis pathways that begin with an NDP-sugar precursor, all synthesis steps occur on the NDP-linkage, that is, the pathway intermediates and the product are always NDP-linked. This makes the QuiNAc synthesis pathway (determined in this study) very unique, that the last step of QuiNAc synthesis (4-reduction) occurs on the Und-PP-linkage after the pathway intermediate is transferred from the UDP-linkage. Although other sugar modifications like O-methylation and O-acetylation may occur even after a polysaccharide is assembled (2,50), the nature of these modifications is different from the 4-reduction step catalyzed by WreQ in QuiNAc synthesis, as they did not change the sugar ring itself.

D. Are there other sugars synthesized in this manner?

Although uniquely determined for QuiNAc synthesis so far, there are other sugars that are potentially synthesized in this manner. It is even possible that sugars shown to be synthesized

entirely on the NDP-linkage *in vitro* may actually be synthesized following a pathway resembling that of QuiNAc *in vivo*. One reason that such pathways were not identified previously is that it would not have been proposed, thus, the biochemical studies were biased toward not testing this possibility. It is worth mentioning that UDP-QuiNAc has been synthesized following the proposed, canonical pathway in this dissertation study, but only after conducting biochemical studies on other enzymes and *in vivo* experiments can we conclude that this is not a bonafide pathway in the cell. One important observation is that WreQ (the 4-reductase) catalyzes the reduction of UDP-ADHexu to UDP-QuiNAc very slowly. In many studies of sugar synthesis, kinetic information was not provided.

Other sugars that may be synthesized in the same manner as QuiNAc are Bac (bacillosamine or 2,4-diamino-2,4,6-trideoxy-D-glucose), AADGal (2-acetamido-4-amino-6-deoxygalactose) and FucNAc (*N*-acetyl-D-fucosamine). All three sugars have the following common features, which are also shared by QuiNAc: First, they have been shown or proposed to be the initiating sugar of a polysaccharide. In other words, they are the first sugar linked to the lipid carrier in the synthesis of a sugar polymer. Bac is the first sugar of the lipid-linked oligosaccharide (LLO) for protein N-glycosylation in *Campylobacter jejuni* (38). AADGal is the first sugar of capsular polysaccharide A (CPSA), which coats the surface of the mammalian symbiont *Bacteroides fragilis* (87). FucNAc is the first sugar of the O-antigen repeating unit in *Pseudomonas aeruginosa* O5 (61). Second, their biosyntheses all begins with the 4,6-dehydration of UDP-GlcNAc. Third, the 4,6-dehydratases involved are integral membrane proteins, which are homologous to WreV (of *R. etli* CE3). The biosynthesis of UDP-Bac and UDP-AADGal has been demonstrated *in vitro* with purified enzymes (38,87). The initiating GTases responsible for *C. jejuni* LLO and *B. fragilis* CPSA synthesis are homologous to WreU of *R. etli* CE3. They have been shown *in vitro* to transfer Bac-1-P or AADGal-1-P to a lipid carrier (65,87). However, the possibility that UDP-ADHexu is a better substrate for these initiating GTases was not investigated. Future studies on initiating GTases of this kind should consider that a pathway

intermediate may be the true substrate. The biosynthesis of FucNAc (D-configuration) has not been studied *in vitro*. The initiating GTase for its transfer in *P. aeruginosa* O5 is $WbpL_{O5}$, a homolog of *P. aeruginosa* O6 $WbpL_{O6}$. The possible FucNAc biosynthesis pathway will be discussed in section E below together with QuiNAc synthesis in *P. aeruginosa* O6.

Sugars that are not the first residue of a polysaccharide structure may also be synthesized in the same manner as QuiNAc. Take QuiNAc as an example, if it is the second (or third, fourth…) sugar of a polysaccharide, it is linked to the non-reducing end of another sugar residue instead of Und-PP. Such sugars are normally transferred from NDP-sugar donors. But it is possible that QuiNAc is still first synthesized as Und-PP-QuiNAc following the determined pathway. In this case, Und-PP-QuiNAc does not serve as the stem for building a polysaccharide chain, but serves as the donor of QuiNAc to another growing polysaccharide. Lipid-linked sugar donors exist in both bacteria and eukaryotes. For example, lipid-linked galacturonic acid (GalA) donates the GalA residue for LPS core modification in *R. leguminosarum* (66). In the biosynthesis of lipid-linked oligosaccharides (LLOs) for protein N-glycosylation in eukaryotes, sugars are donated by lipid-linked sugar donors during the synthesis phase in the ER lumen (86). However, it should be pointed out that these lipid-linked sugar donors are different from Und-PP-QuiNAc in two major ways. First of all, the sugar moiety of the lipid-linked sugar donors is transferred from an NDP-sugar. In this transfer reaction, only the sugar is transferred to the lipid carrier, resulting in a lipid-phosphate-sugar (lipid-P-sugar), whereas Und-PP-QuiNAc contains a pyrophosphate linkage; Secondly, the sugar moieties of the lipid-linked sugar donors are in their "final" form and are not further modified after transferred to the lipid, whereas the QuiNAc of Und-PP-QuiNAc is modified from ADHexu on the Und-PP linkage. Given these differences, whether Und-PP-QuiNAc or another lipid-PP-sugar can serve as a sugar donor requires further biochemical study.

E. Biosynthesis of QuiNAc and FucNAc in bacteria with PNPT family initiating GTases

QuiNAc has been identified and proposed to be the first sugar transferred onto a lipid carrier in bacteria that contain PNPT type initiating GTase. It is unclear whether QuiNAc biosynthesis in these bacteria follows the same pathway as determined in *R. etli* CE3, which has a PHPT type initiating GTase (WreU). FucNAc is the C-4 epimer of QuiNAc, its synthesis is expect to differ from QuiNAc only by the last step, catalyzed by a 4-reductase that has a stereospecificity for *galacto*- configuration (instead of a *gluco*- configuration for QuiNAc). Thus, FucNAc biosynthesis should follow in the same manner as QuiNAc.

P. aeruginosa serotype O5 and O6 are the best candidates for such investigation. FucNAc and QuiNAc have been proposed to be the initiating sugar of the repeating unit of the serotype-specific B-band O antigen in O5 and O6, respectively (43,61). Genetic studies of O-antigen synthesis genes suggested that WbpL was the initiating GTase. The $wbpL_{O5}$ and $wbpL_{O6}$ mutants did not make B-band O antigen and the *wbpL* gene from the two serotypes can cross-complement each other (43). Based on these results, *wbpL* was postulated to have relaxed specificity, being able to use both UDP-QuiNAc and its C-4 epimer UDP-FucNAc as substrate.

However, the activity of the *R. etli* CE3 initiating GTase WreU determined in this dissertation inspired another possible explanation: $WbpL_{O5}$ and $WbpL_{O6}$ may both use the 4-keto intermediate UDP-ADHexu as substrate. Thus, $wbpL_{O5}$ and $wbpL_{O6}$ can definitely complement each other. The reduction to FucNAc or QuiNAc is catalyzed by stereospecific 4-reductases on the Und-PP linkage. In this scenario, the *P. aeruginosa* initiating GTase WbpL catalyzes the same reaction as WreU of *R. etli* CE3, and the biosynthesis of FucNAc and QuiNAc in *P. aeruginosa* follows the same pathway as in *R. etli* CE3.

The key to this argument is whether WbpL can use UDP-ADHexu as a substrate. One line of evidence suggests that a WbpL homolog may be able to use UDP-ADHexu as substrate, which is discussed below. The 4-keto intermediate in the QuiNAc synthesis, ADHexu, has been

found to naturally exist in polysaccharides of several bacterial strains (40,88-90). In *Yersinia enterocolitica* serotype O:3 wild-type strain, ADHexu was proposed to be the first sugar of the outer core region of LPS (equivalent to the non-repeating portion of *R. etli* CE3 O antigen) (40). Genetic studies suggested that the *wbcO* gene encoded the initiating GTase and proposed that it used UDP-ADHexu as substrate (91). WbcO belongs to the PNPT family and it was classified in the same subgroup as WbpL (92,93). One simple experiment to test whether WbpL can use UDP-ADHexu as substrate is to complement a *Y. enterocolitica* O:3 *wbcO* mutant with the *wbpL* gene. However, the best way to determine the substrate specificity of WbpL (and WbcO) is to carry out reactions *in vitro* with UDP-ADHexu or UDP-QuiNAc as substrate. Another intriguing test is to complement an *R. etli wreU* mutant with *wbcO* (or *wbpL*). The result of this test may provide implications beyond just enzymatic activity, but regarding interaction with other enzymes.

F. The possibility of a membrane-located complex formed by enzymes involved in QuiNAc biosynthesis

In the model presented in Fig. 37B, *R. etli* CE3 WreU, WreV and WreQ proteins were proposed to form a membrane-bound complex. WreV belongs to a subgroup of UDP-GlcNAc 4,6-dehydratases that are intergral membrane proteins. Because NDP-sugars are water-soluble compounds, it seems unnecessary for a NDP-sugar 4,6-dehydratase to be on the membrane. Indeed, most other 4,6-dehydratases are soluble proteins. This suggests that WreV and its homologs are located on the membrane for a reason. Initiating GTases, such as WreU (*R. etli*) and WbpL (*P. aeruginosa*) are integral membrane proteins. In *R. etli* CE3, WreU uses the UDP-ADHexu (the product of WreV) as substrate for catalyzing the transfer reaction. It may directly acquire UDP-ADHexu from WreV through interacting with it. The 4-reductase WreQ may interact with WreU and reduce the ADHexu moiety after it is transferred onto the lipid carrier. The organization of the three enzymes into a complex may not only make QuiNAc synthesis and O-antigen initiation more efficient, but also may be protective. Without complex formation,

UDP-ADHexu synthesized by WreV may be released into the cytosol, where it can encounter the reductase WreQ and be reduced to UDP-QuiNAc. As discussed earlier, UDP-QuiNAc is potentially toxic to the cell. Also, UDP-QuiNAc is not a preferred substrate of the initiating GTase WreU.

Attempts to express WreU and WreV separately as His-tagged proteins in *E. coli* and purify them were not successful. WreU could be expressed but it may aggregate on the membrane or form inclusion bodies, which prevent its purification. WreV expression cannot be detected in *E. coli*; possibly, the WreV protein is degraded very soon after expression. If the two proteins naturally form a complex in *R. etli* CE3, they may require co-expression for stability. Thus, expressing them separately in *E. coli* may have suffered from the absence of their partner. Co-expression may solve the problem and improve the expression and solubilities of both proteins.

If QuiNAc and FucNAc are produced in the same manner in *P. aeruginosa* as QuiNAc produced in *R. etli* CE3, then the initiating GTase WbpL, the 4,6-dehydratase WbpM and the 4-reductase WbpV (O6) or WbpK (O5) may also form a complex. Because WbpL is structurally quite different from WreU, the organization of the complex may also be different from the one in *R. etli* CE3.

G. The location of QuiNAc biosynthetic genes in *Rhizobium* species and evolutionary implications

As determined in Chapter Four, all sequenced *R. etli* and *R. leguminosarum* strains contain homologs of *R. etli* CE3 *wreU*, *wreV* and *wreQ* gene. However, the large chromosomal O-antigen genetic cluster of *R. etli* CE3 is not conserved. O-antigen structures determined from *Rhizobium* strains vary greatly (4,67,69), even though all of them contain QuiNAc. It can be speculated that the ability to synthesize and attach QuiNAc was acquired separately from the rest of the O-antigen genes. This acquisition occured prior to the time these species diverged from closely related lineages. Thus, QuiNAc very likely is the common primer sugar among all *R. etli*

strains and *R. leguminosarum* strains. The O-antigen genetic clusters separately acquired by these strains accommodated this primer and resulted in varying O-antigen structures.

The *wreQ* gene is separated from the *wreU* and *wreV* genes. It is unclear whether they were acquired separately, or whether they were acquired together and then separated. The fact that *wreU* and *wreV* are in close proximity makes sense because the two genes are sufficient for an initiating reaction, resulting in the ADHexu on the Und-PP linkage. It is possible that ADHexu was once a component of the O antigen or another polysaccharide. In that case, the replacement of it by QuiNAc was a later event following the introduction of the *wreQ* gene. Note that such an event is dependent on the degree of specificity of the succeeding enzyme in the synthesis of the O antigen (or other polysaccharide) if the polysaccharide remains the same otherwise. This thinking leads to the question of whether *wreQ* introduction might be coordinated with the introduction of other genes.

H. Potential applications of the knowledge of QuiNAc biosynthesis

Since many of the LPS monosaccharide components are rare sugars and only present in certain pathogenic bacterial species, these unusual sugars and the enzymes involved in their synthesis can be targets for novel antimicrobial drug development. An in-depth understanding of the biosynthetic pathways of these sugars and the mechanisms of the encoded enzymes is an essential first step to undertake.

The primary role of O antigens appears to be protective and they are known as essential virulence determinants of pathogenic bacteria (2). So far, the sugar QuiNAc has only been reported in bacteria. This makes it a potential target for the treatment of bacterial infections in eukaryotes. Since QuiNAc is usually the primer sugar in O antigens, disrupting its biosynthesis may lead to the total loss of O antigen and reduced virulence. Inhibiting the enzymes required for QuiNAc synthesis is less likely to disrupt eukaryotic cell functions as eukaryotes do not make

QuiNAc and may not contain a homolog of these enzymes. One potential target is the initiating GTase, especially the PHPT family enzymes, as they have only been found in bacteria while members of the PNPT family are involved in N-glycosylation in eukaryotes (93).

CHAPTER SIX: METHODS

A. Bacterial Strains and Growth Conditions

Rhizoubium etli CE3 was derived from *R. etli* type strain CFN42 by a spontaneous mutation conferring resistance to streptomycin (94). As in all past studies of *wreQ* and almost all other studies of the LPS of *R. etli* CFN42, strain CE3 was the wild-type source of DNA and genotype for strain constructions. All *R. etli* strains were grown to stationary phase at 30 °C in TY liquid medium (0.5% tryptone (Difco Laboratories), 0.3% yeast extract (Difco), and 10 mM $CaCl_2$). Unless stated otherwise, all *Escherichia coli* strains were grown to stationary phase at 37 °C in Luria-Bertani (LB) liquid medium (1.0% tryptone, 0.5% yeast extract, and 0.5% NaCl). Agar medium contained 1.5% Bacto Agar (Difco).

B. Computer Analysis of Predicted Protein Sequences

Amino acid sequences of open reading frames (ORFs) were downloaded from National Center for Biotechnology Information (NCBI) database. BLAST searches were performed against database Non-redundant protein sequences (nr) using Blastp (protein-protein BLAST) algorithm (95). Alignment of predicted protein sequences was performed using Clustal Omega at the European Bioinformatics Institute (EBI) website (http://www.ebi.ac.uk/Tools/msa/clustalo/). Predicted transmembrane domains were determined by Transmembrane Helix Prediction (TMHMM) version 2.0 at the Center of Biological Sequence Analysis (CBS) at the Technical University of Denmark website (http://www.cbs.dtu.dk/services/TMHMM/).

C. DNA Techniques

R. etli genomic DNA was isolated using a GenElute Bacterial genomic DNA kit (Sigma-Aldrich). Plasmid DNA isolation from *E. coli* cultures and purification from agarose gels was

performed using the QIAprep Spin Miniprep Kit (Qiagen) and Gel/PCR DNA Fragments Extraction Kit (IBI Scientific), respectively. Amounts of DNA fragments were amplified using the Expand High Fidelity PCR System (Roche Applied Science) with custom primers (Eurofins MWG Operon) containing restriction sites to facilitate cloning. Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (NEB). Vectors or recombinant plasmids were transformed into INVαF', OmniMAX (Invitrogen) or 5-alpha (NEB) *E. coli* cells for cloning or into BL21(DE3) *E. coli* cells (Lucigen) for protein overexpression. DNA sequencing was performed by Functional Biosciences.

D. Site-directed mutagenesis

The general approach of site-directed mutagenesis is to first create a mutated copy of a gene on a plasmid, then transfer this plasmid into the bacterial cell to replace the wildtype copy of the same gene by homologous recombination, and finally the double recombinant was selected. The specific cases for each gene are described below.

R. etli CE3 wreV gene— the *wreV* gene with about 500 bp flanking sequence on each side was amplified from CE3 genomic DNA by polymerase chain reaction (PCR) using primers wreVf-EcoRI (5'-GCCGAATTCTCCAGTCTCGACGAAC-3') and wreVr-SphI (5'-GCCGCATGCAAGCACCAATGGTGAT-3'). The 2928 bp PCR product was digested with restriction enzymes EcoRI and SphI and then ligated to pEX18Tc vector (96), creating plasmid pTL2. pTL2 was digested with PstI, which deleted a 279 bp fragment from *wreV*, and the Gm cassette from plasmid pUCGm (97) or the Km cassette from plasmid pBSL86 (98) was inserted into the PstI site, creating plasmid pTL4 and pTL3, respectively.

R. etli CE3 wreQ gene— ~1700 bp SacI-SphI fragment containing the *wreQ* gene was cut from plasmid pJBQ1 (19), then ligated to plasmid pEX18Tc, creating plasmind pTL9. pTL9

was digested with SalI, which deleted a 654 bp fragment from *wreQ*, and the Gm or Km cassette was inserted, creating plasmid pTL10 ad pTL11, respectively.

R. etli CE3 O-antigen ligase gene on plasmid B (oal_{pB})— Because there are no appropriate restriction sites for generating a mutant copy of this gene, a three-piece-ligation approach was taken (Fig. 41). The initial 345 bp of the oal_{pB} gene together with upstream sequences was amplified by PCR using primers upforwardOAL (5'-GCGAATTCCGATGCTGGTCGAGACT-3') and upreverseOAL (5'-GCGGTACCAGGCGGATTGACGAATA-3'). The last 16 bp of the oal_{pB} gene together with downstream sequences was amplified by PCR using primers DwnForwardOAL (5'-GCGGTACCTGTGGTTCGAAGCTGAT-3') and DwnReverseOAL (5'-GCGCATGCTGAAGCTGGCCGAAGTT-3'). The PCR products were ligated to pCR2.1 vector (Invitrogen), creating plasmids pMM1 and pMM2, respectively. pMM1 was digested with EcoRI and KpnI to release a ~1.2 kb fragment . pMM2 was digested with KpnI and SphI to release a ~0.9 kb fragment. The two fragments were ligated with pEX18Tc vector (three-piece-ligation), creating plasmid pLS35. pLS35 was digested with KpnI and the Gm cassette from plasmid pUCGm was ligated into this site to create plasmid pLS36.

R. etli CE3 O-antigen ligase gene on the chromosome (oal_{chr})— this gene copy was also mutated with the three-piece-ligation method as described above. The initial 51 bp of the oal_{chr} gene together with upstream sequences was amplified by PCR using primers oal(c)upfor (5'-GCGAATTCCTTGATCCGAGTGCACC-3') and oal(c)uprev (5'-GCGAGCTCAGGCGGATTGATGAAGA-3'). The last 26 bp of the oal_{pB} gene together with downstream sequences was amplified by PCR using primers oal(c)dwnfor (5'-GAGAGCTCGATTTGCGCCCGTCGTT-3') and oal(c)dwnrev (5'-GAGCATGCCGCCATGGTGACGACTG-3'). The PCR products were ligated to pCR2.1 vector (Invitrogen), creating plasmids pLS38 and pLS39, respectively. pLS38 was digested with EcoRI and SacI to release a ~1000 bp fragment . pLS39 was digested with SacI and SphI to release a

~960 bp fragment. The two fragments were ligated with pEX18Tc vector (three-piece-ligation), creating plasmid pLS40. pLS40 was digested with SacI and the Gm cassette from plasmid pUCGm, or the Km cassette from plasmid pBSL86 was ligated into this site, creating plasmids pLS41 and pLS42.

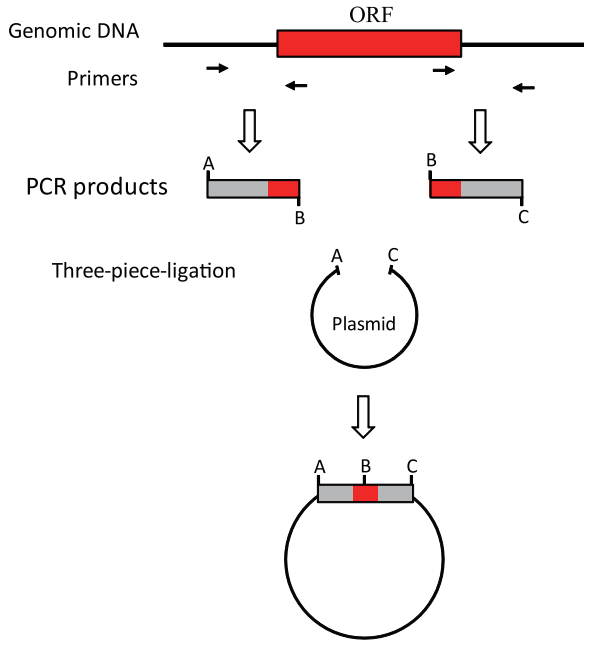


Figure 41. Schematic diagram showing the "three-piece-ligation" method for mutagenesis. The 5' region of an ORF with upstream sequence and the 3' region of an ORF with downstream sequence were separately amplified by PCR. The two PCR products and a plasmid were ligated via designed restriction sites (A, B and C) on the primers. The resulting plasmid carries the ORF with a deletion and the up- and downstream region for homologous recombination. An antibiotic cassette can then be inserted in the mutated ORF through restriction site B (not shown).

R. leguminosarum bv. viciae 3841 wreU gene homolog (wreU$_{3841}$)— wreU$_{3841}$ was amplified from *R. leguminosarum* bv. *viciae* 3841 genomic DNA as a 2719 bp fragment using primers 3841wreUmutF (5'-GCGAATTCCTCGACCGATGACGTGA-3') and 3841wreUmutR (5'-GCGGTACCAACGTGCGAACATTGAC-3'). The PCR product was ligated to plasmid pCR2.1, creating plasmid pTL12. pTL12 was digested with EcoRI and KpnI, and the resulting fragment was ligated to plasmid pEX18Tc, creating plasmid pTL13. The Km cassette from pBSL86 was inserted in the SalI site of pTL13, creating plasmid pTL15.

R. leguminosarum bv. viciae 3841 wreQ gene homolog (wreQ$_{3841}$)— wreQ$_{3841}$ was amplified from 3841 genomic DNA as a 2941 bp fragment using primers 3841wreQmutF (5'-GCTCTAGACCGTTCTGCTTGCCACC-3') and 3841wreQmutR (5'-GCGCATGCAACTGGAAGGCCTATCG-3'). The PCR product was ligated to plasmid pCR2.1, creating plasmid pLS30. pLS30 was digested with XbaI and SphI to release the fragment containing *wreQ$_{3841}$*, which was ligated to plasmid pEX18Tc, creating plasmid pLS32. pLS32 was digested with PstI and the Km cassette from plasmid pBSL86 was ligated into this site, creating plasmid pLS33.

Replacement of wild-type locus by homologous recombination—Plasmids carrying the mutated copy of genes were transferred into *R. etli* CE3 (or *R. leguminosarum* bv. *viciae* 3841) by triparental mating (99) with mobilizing-plasmid containing strain MT616 (100). Specifically, 800 µl CE3, 200 µl MT616 and 200 µl *E. coli* strain harboring the mutant construct were taken from overnight cultures and mixed on TY agar plates. After overnight incubation at 30 °C, colonies on the plate were resuspended in 2 ml liquid TY media and transconjugates containing the mutant constructs were selected on two types of TY agar plates: (1) supplemented with 200 µg of streptomycin/ml, 30 µg of nalidixic acid/ml, and either 30 µg of gentamycin/ml or 30 µg of kanamycin/ml (depending on the antibiotic cassette inserted in the mutated gene copy), and (2) plates supplemented with 8 % sucrose in addition to the above antibiotics (the rationale for this selection is explained in (99)). Colonies from the sucrose-containing plate were streaked on the two kinds of plates above and TY agar plates supplemented with 1 µg of tetracycline/ml. The tetracycline-sensitive, sucrose-resistant colonies were saved after verification of the double recombination event by PCR. For selecting mutants of *R. leguminosarum* strains, the antibiotic nalidixic acid was omitted from the corresponding plates for *R. etli* strains. Following this method, plasmids pTL4 (*wreV*::Gm), pTL3 (*wreV*::Km), pTL11 (*wreQ*::Km), pLS36 (*oal$_{pB}$*::Gm) or pLS41 (*oal$_{chr}$*::Gm) was transferred into *R. etli* CE3. The resulting *R. etli* strains were CE568, CE569, CE572, CE575 and CE577, respectively. Plasmids pLS33 (*wreQ$_{3841}$*::Km) or pTL15

(*wreU*$_{3841}$::Km) was transferred into *R. leguminosarum* bv. *viciae* 3841, resulting in strain LT9 and LT11.

Construction of R. etli O-antigen ligase double mutant— plasmid pLS42 (*oal*$_{chr}$::Km) was transferred into *R. etli* strain CE575 (*oal*$_{pB}$::Gm) by triparental mating. The resulting *R. etli* strain was CE578 (*oal*$_{pB}$::Gm, *oal*$_{chr}$::Km).

E. Complementation of mutants

For complementation of *R. etli* (or *R. leguminosarum*) mutants, complementing DNA fragments were cloned into stable RK2-derived cloning vector pFAJ1708 (51), plasmid pJB21 (derived from pFAJ1700), or plasmid pJBC1 (101). The resulting constructs were transferred into mutant strains by triparental mating.

● Complementation of mutants with the corresponding wildtype genes

Complementation of R. etli wreV mutants—the *R. etli* CE3 *wreV* gene with flanking sequence was amplified from CE3 genomic DNA by PCR using primers wreVf-EcoRI (5'-GCCGAATTCTCCAGTCTCGACGAAC-3') and wreVcomp (5'-GCCGGTACCTATCGCGTCCGTGAAC-3'). The 2969 bp PCR product was ligated to plasmid pCR2.1 (Invitrogen), creating plasmid pTL6. pTL6 was digested with EcoRI and KpnI and the resulting fragment carrying *wreV* was ligated to plasmid pFAJ1708 or pJB21 (J. Box, unpublished), creating plasmids pTL7 and pTL8, respectively. (Only the result of complementation with pTL7 was shown in this dissertation, but pTL8 gave the same result.)

Complementation of R. etli wreQ mutant— plasmid pJBQ3 was used for this complementation. pJBQ3 is a pFAJ1708 derivative carrying a DNA fragment containing the *R. etli wreQ* gene(19).

● Complementation of mutants with homologous genes from other bacterial species or other strains of the same species

Complementation of R. etli wreV mutants with P. aeruginosa wbpM— the *wbpM* gene was amplified from the pBAD-*wbpM* vector (provided by Dr. J. S. Lam) using primers pBADwbpMforward (5'-GGCATATGTTGGATAATTTGAGG-3') and pBADwbpMreverse (5'-TTAAGCTTTCAGGGTTCTCGCCGCC-3'). The PCR product was ligated to plasmid pCR2.1. This construct was digested with NdeI and HindIII, and the resulting fragment was ligated to plasmid pJBC1 (102), generating plasmid pTL61.

Complementation of the R. etli wreQ mutant with P. aeruginosa wbpV— plasmid pFV611-26 (43) (provided by Dr. J. S. Lam, University of Guelph, Guelph, Canada) was digested with PstI and SacI to release a fragment containing the *wbpV* gene of *P. aeruginosa* O6. This fragment was inserted into two pFAJ1700 based vectors (51), generating pJB7 and pJB8, respectively.

Complementation of R. leguminosarum wreQ, wreU mutants with corresponding genes of R. etli CE3— plasmid pJBQ3 (described above) was used for complementation of *R. leguminosarum wreQ* mutant LT9. Plasmid pKT100 was used for complementation of *R. leguminosarum wreU* mutant LT11. pKT100 is a pJB21 derivative carrying the *R. etli wreU* gene (Kristylea Ojeda, unpublished).

● Transfer of constructs into mutant strains

pTL7, pTL8 and pTL61 were separately transferred into CE568 and CE569; pJBQ3 was transferred into CE572 and LT9; pJB7 and pJB8 were transferred into CE166; pKT100 was transferred into LT11 by triparental mating with mobilizing plasmid-containing strain MT616. Strain containing these constructs were selected on TY agar plates supplemented with 200 μg of streptomycin/ml, 30 μg of nalidixic acid/ml (omit this antibiotic when selecting *R. leguminosarum* strains), 5 μg tetracycline/ml, 30 μg of kanamycin/ml or 30 μg of gentamycin/ml (depending on the recipient strain). Single colonies were purified and analyzed by SDS-PAGE.

F. Analysis of LPS by SDS-PAGE and LPS sugar compositions

LPS was prepared from 0.5 ml of full-grown cultures and analyzed by SDS-PAGE with periodate-silver stain (21). To determine whether QuiNAc was present, LPS was extracted from washed bacterial pellets (from 4.0 L of culture) by the hot phenol-water method, dialyzed, treated with RNase A, DNase I and protease K, dialyzed again followed by lyophilization (21). For determination of neutral and amino sugars, LPS preparations were treated with 2 M trifluoroacetic acid for 2 h at 121°C. After reduction with $NaBD_4$ and acetylation, alditol acetate derivatives of the LPS sugars were analyzed by gas chromatography (GC) under conditions described previously (21).

G. Constructs for protein overexpression

Construct for His_6-WreQ overexpression— R. etli CE3 *wreQ* was amplified from the genomic DNA of *R. etli* CE3 using primers lpsQ-over1 (5'-CCGGCATATGCGATGCCTCGTCACC-3') and lpsQ-over2 (5'-GGCCGGATCCCTATTCCACTGCAAG-3'). The ~950 bp PCR product was ligated to plasmid pCR2.1, creating plasmid pLS2. pLS2 was digested with NdeI and BamHI, the *wreQ*-containing fragment was inserted into the pET15b vector (Novagen), generating plasmid pLS7. The construct was checked by restriction digestion and nucleotide sequence determination. It encodes the WreQ protein with additional amino acids at the amino-terminus, MGSSHHHHHHSSGLVPRGSH (the His_6 tag is underlined). The WreQ construct is referred to as His_6-WreQ in this work.

Construct for His_6-WreU overexpression— R. etli CE3 *wreU* gene was amplified from the genomic DNA of *R. etli* CE3 using primers Lpsβ1-over1 (5'-CCGGCATATGGGCTTGAAACGGGCG-3') and Lpsβ1-over2 (5'-GGCCGGATCCCTAGTGCTTTATTCC-3'). The PCR ~590 bp product was ligated into plasmid

pCR2.1, creating plasmid pKP2. pKP2 was digested with NdeI and BamHI and the *wreU*-containing fragment was inserted into pET15b vector (Novagen), generating plasmid pLS5. The construct was checked by restriction digestion and nucleotide sequence determination. It encodes the WreU protein with additional amino acids at the amino-terminus, MGSSHHHHHHSSGLVPRGSH (the His_6 tag is underlined). This WreU construct is referred to as His_6-WreU in this work.

Construct for WreG-His_6 overexpression— R. etli CE3 *wreG* gene was amplified from the genomic DNA of *R. etli* CE3 using primers pET-wreGfor (5'-GCGCTAGCATGAGAGTCCTTCATTT -3') and wreGrevXhoI (5'-TTCTCGAGGCGGGAACCGGCCACGT-3'). The ~1130 bp PCR product was digested with NheI and BamHI and the *wreG*-containing fragment was inserted into pET21b vector (Novagen), generating plasmid pTL59. The construct was checked by restriction digestion and nucleotide sequence determination. It encodes the WreG protein with amino-terminal additional amino acids MAS, and carboxy-terminal additional amino acids LEHHHHHH (the His_6 tag is underlined). This WreG construct is referred to as WreG-His_6 in this work.

Transform constructs into E. coli cells for overexpression— plasmid pLS7, pLS5 and pTL59 were separately transformed into *E. coli* BL21(DE3) cells by electroporation.

H. Test the functions of the His-tagged constructs *in vivo* by complementation

Complementation of the R. etli wreQ mutant with His_6-wreQ— using the constructed plasmid pLS7 as template, the *His_6-wreQ* gene together with the vector ribosome binding site was amplified with primers UQHiscomp-up (5'-GCCGAATTCATACCCACGCCGAAACAAG-3') and UQHiscomp-dwn (5'-GCCGGTACCAGTTCCTCCTTTCAGCAAA-3'). The ~1390bp PCR product was ligated into plasmid pCR2.1, creating plasmid pLS19. pLS19 was digested with

EcoRI and KpnI, the *His$_6$-wreQ*-containing fragment was ligated into plasmid pFAJ1708 (51), generating plasmid pLS31.

Complementation of R. etli wreU mutant with His$_6$-wreU—the *His$_6$-wreU* gene together with the vector ribosome binding site was amplified from the constructed plasmid pLS5, with primers UQHiscomp-up (5'-GCCGAATTCATACCCACGCCGAAACAAG-3') and UQHiscomp-dwn (5'-GCCGGTACCAGTTCCTCCTTTCAGCAAA-3'). The ~1020 bp PCR product was ligated into plasmid pCR2.1, creating plasmid pLS18. pLS18 was digested with EcoRI and KpnI and the *His$_6$-wreU*-containing fragment was ligated into plasmid pFAJ1708 (51), generating plasmid pLS22.

Complementation of R. etli wreG mutant with wreG-His$_6$—using the constructed plasmid pTL59 as template, the *wreG-His$_6$* gene together with the vector ribosome binding site was amplified with primers UQHiscomp-up (5'-GCCGAATTCATACCCACGCCGAAACAAG-3') and UQHiscomp-dwn (5'-GCCGGTACCAGTTCCTCCTTTCAGCAAA-3'). This ~1570 bp PCR product was ligated into plasmid pCR2.1, creating plasmid pTL62. pTL62 was digested with XbaI and KpnI and the *wreG-His$_6$* containing fragment was ligated into plasmid pFAJ1708 (51), generating plasmid pTL63.

Transfer of his-tag complementation constructs into R. etli mutant strains— pLS31 (*His$_6$-wreQ*) was transferred into CE166 (*wreQ*::Tn*5*); pLS22 (*His$_6$-wreU*) was transferred into CE566 (*wreU*::Km); pTL63 (*wreG-His$_6$*) was transferred into CE358 (*wreG*::Tn*5*) by triparental mating with mobilizing plasmid-containing strain MT616. Strain containing these constructs were selected and tested by SDS-PAGE as described above.

I. Protein overexpression

Overexpression of WbpM-His-S262— The *E. coli* BL21(DE3)plysS strain carrying a pET vector derivative encoding a soluble truncated version of the WbpM protein (residue 262-665)

(37) was provided by Dr. J. S. Lam. To express the WbpM-His-S262 protein, one liter LB was inoculated with 20 ml overnight culture of the BL21(DE3)plysS strain and grown at 37 °C. Protein expression was induced by addition of 0.5 mM IPTG. After 5 h induction at 37 °C, the cells were harvested by centrifugation at 5000 g for 15 min at 4 °C, and the pellets was stored at -80 °C until needed.

His_6-WreQ Overexpression— BL21(DE3)/pLS7 strain grown overnight in 5 ml LB medium containing ampicillin (100 μg/ml) was used to inoculate a flask of one liter LB medium. The culture was grown at 37 °C with shaking until OD_{600} reached 0.6. His_6-WreQ expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) (Gold Biotechnology) to a final concentration of 0.1 mM and the culture was shaken for a further 20 hours at 16 °C. Cells were harvested by centrifugation at 5000 g for 15 min at 4 °C, and the pellets was stored at -80 °C until needed.

Overexpression of DGK, His_6-WreU and WreG-His_6— E. *coli* BL21 Gold cells carrying a pET vector construct encoding the *Streptococcus mutans* diacylglycerol kinase (DGK) with a carboxy-terminal His_6-tag (62) was provided by Dr. Barbara Imperiali, Massachusetts Institute of Technology, Cambridge, MA. His_6-WreU and WreG-His_6 were expressed from BL21(DE3)/pLS5 and BL21(DE3)/pTL59 strains, respectively.

Expression of DGK, His_6-WreU and WreG-His_6 followed the same procedure: A flask of one liter LB medium containing appropriate antibiotics [ampicillin (100 μg/ml) for His_6-WreU and WreG-His_6, kanamycin (50 μg/ml) for DGK] was inoculated with a 5 ml overnight start culture and shake at 37 °C until an optical density between 0.6 and 0.8 was reached. The flask was covered with ice and chilled for 1 hour. Protein expression was induced by adding IPTG to the culture (final IPTG concentration: 1 mM for DGK, 0.01 mM for His_6-WreU and 0.1 mM for WreG-His_6), and the culture was shaken for a further 20 hours at 16 °C. Cells were harvested by centrifugation at 5000 g for 15 min at 4 °C, and the pellets was stored at -80 °C until needed.

J. Preparation of Membrane Fractions

Membrane fractions were prepared from *E. coli* cells expressing DGK, His_6-WreU and WreG-His_6 for use as crude enzyme or for purification of membrane located proteins.

DGK—The frozen cell pellets from 500 ml culture were thawed into 15 ml of buffer E (50 mM Tris, 1 mM ethylenediaminetetraacetic acid) and lysed by sonication. A low speed spin (6000 g, 20 min at 4 °C) removed most of the cellular debris; it was followed by high speed spin (65,000 g, 120 min at 4 °C) to pellet the cell membrane fraction. The pellet was preserved at -80 °C.

His_6-WreU—The frozen cell pellets from 500 ml culture were thawed into 12 ml of buffer F (20 mM Tris, 300 mM NaCl, pH 8.5) and lysed by sonication. Cellular debris was removed by centrifugation (10,000 g for 15 min at 4 °C), and the supernatant was centrifuged at 65,000 g for 120 min at 4 °C to pellet the cell membrane fraction. The pellet was homogenized in 0.5 ml buffer F and aliquoted into 100 µl fractions for storage at -80 °C.

WreG-His_6—The frozen cell pellets from 500 ml culture were thawed into 12 ml of buffer A (20 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, pH=7.0, and 14.3 mM 2-Mercaptoethanol) and lysed by sonication. Cellular debris was removed by centrifugation (10,000 g for 15 min at 4 °C), and the supernatant was centrifuged at 65,000 g for 120 min at 4 °C to pellet the cell membrane fraction. The pellet was homogenized in 1 ml buffer A and aliquoted into two 0.5 ml fractions before preserved at -80 °C.

K. Protein purification

Purification of His_6-WreQ— The His_6-WreQ protein was purified with Ni^{2+} affinity chromatography. Specifically, the frozen cell pellets were resuspended in buffer A (20 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, pH=7.0, and 14.3 mM 2-Mercaptoethanol) and disrupted by sonication. Cellular debris was removed by centrifugation (6000 g for 20 min at

4 °C), and the supernatant was centrifuged at 65,000 g for 120 min at 4 °C. The supernatant was passed through a column containing 1 ml Ni^{2+}-profinity IMAC resin (Bio-Rad). The column was washed with 10 ml buffer A followed by washing with buffer B (20 mM sodium phosphate, 300 mM NaCl, 20 mM imidazole, pH=7.0, and 14.3 mM 2-Mercaptoethanol). His_6-WreQ protein was eluted with buffer C (20 mM sodium phosphate, 300 mM NaCl, 300 mM imidazole, pH=7.0, and 14.3 mM 2-Mercaptoethanol). Triton X-100 was added to the combined elution fraction to a final concentration of 0.1 %. Then, the elution fraction was transferred to a dialysis tubing (molecular weight cut off: 12-14,000 Da) and dialyzed overnight at 4 °C in 1 liter buffer containing 20 mM sodium phosphate, 300 mM NaCl, 14.3 mM 2-Mercaptoethanol, pH=7.0. The dialyzed protein sample was concentrated with Amicon Ultra-15 (nominal molecular weight limit: 10 kDa) filter device. The concentrated protein was fast-frozen on dry ice and preserved at -80 °C until needed. Typical yield of purified protein was 4.5 mg per liter of cell culture. Protein concentration was determined by the method of Bradford (103).

Purification of WbpM-His-S262— Purification of WbpM-His-S262 was carried out in the same manner as described above for His_6-WreQ, except that all the buffers (buffer A, B and C) do not contain 2-Mercaptoethanol. Following purification, the elution fraction was dialyzed against 1 L buffer (20 mM sodium phosphate, 300 mM NaCl, pH=7.0) and concentrated with Amicon Ultra-15 (nominal molecular weight limit: 10 kDa) filter device. Protein concentration was determined by the method of Bradford (103).

Purification of DGK from membrane fractions— Cell membrane pellet was thawed and resuspended in 0.5 ml buffer G (20 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, pH 8.0) and incubated with 1% CHAPS for 1 hour at 4 °C to solubilize membrane proteins. Then the sample was incubated with 250 µl Ni^{2+}-profinity IMAC resin (Bio-Rad) for 30 min at 4 °C. The resin was placed in a 0.2 µm filter in a microcentrifuge tube for the subsequent wash and elution steps. The resin was washed twice with 375 µl of buffer G containing 1% CHAPS, and twice with 375 µl of the same buffer with 45 mM imidazole. The protein was eluted three times in 200 µl of

the same buffer containing 300 mM imidazole. The combined elution fraction was dialyzed and concentrated with Amicon Ultra-0.5 (nominal molecular weight limit: 10 kDa) filter device. The concentrated protein was aliquoted into smaller fractions for storage at -80 °C.

Purification of WreG-His$_6$ from membrane fractions—Thaw one tube of 0.5 ml cell membrane fraction and incubated with 1% Triton X-100 for 2.5 hour at 4 °C. Then the sample was incubated with 200 µl Ni^{2+}-profinity IMAC resin (Bio-Rad) for 30 min at 4 °C. The resin was placed in a 0.2 µm filter in a microcentrifuge tube for the subsequent wash and elution steps. The resin was washed twice with 250 µl of buffer A containing 0.1% Triton X-100, and twice with 250 µl of the same buffer with 20 mM imidazole. The protein was eluted twice in 250 µl of the same buffer containing 300 mM imidazole. Protein sample dialysis and concentration were performed exactly as described for DGK above.

L. Enzyme assay for WbpM and WreQ

WbpM enzyme assay— Enzyme assays were carried out in 10-µl reactions containing buffer D (20 mM sodium phosphate, pH=7.0), 0.5 mM UDP-GlcNAc (Sigma), 0.029 µM ^{3}H-labeled UDP-GlcNAc (0.1 mCi/mmol PerkinElmer Life Sciences) and 9 µg WbpM. The reaction was incubated at 30 °C and allowed to proceed for 30 min for complete substrate conversion.

WbpM-WreQ coupled enzyme assay— The WreQ catalyzed reaction was carried out by adding 15 µg His$_6$-WreQ protein and 1 mM NAD(P)H to the WbpM reaction described above and allowed to proceed for another 4 hours. In the control reactions where one or multiple components were omitted, the lost volume was compensated with buffer D.

To estimate the rate of the WreQ reaction, 10 µg His$_6$-WreQ protein and 1 mM NAD(P)H were added to the WbpM reaction and the WreQ reactions were quenched at 1 min, 5 min and 30 min.

Separation of sugar nucleotides by thin layer chromatography— The WbpM reaction or the WbpM-WreQ coupled enzyme reaction was stopped by adding an equal volume of ice-cold 100% ethanol and proteins were removed by centrifugation (16,000 g 10 min, 4 °C) after overnight storage at -20 °C. The supernatant was lyophilized to dryness. The dried residues of radioactive reactions were dissolved in 5 µl glass distilled water. The entire volume of each reaction sample was spotted on an aluminum-backed precoated Silica gel 60 plate (EMD Chemicals) and developed in Solvent A (2-propanol/ammonium hydroxide/water, 6:3:1). Dried TLC plates were sprayed with EN^3HANCE Autoradiography Enhancer (PerkinElmer Life Sciences) and exposed to film (Kodak BioMax XAR Film) for 3 days at -80 °C.

M. GC-MS analysis of glycosyl residue

To detect the glycosyl residue of the final product of the WbpM-WreQ reaction, the reaction was carried out in a volume of 200 µl containing buffer D, 1 mM UDP-GlcNAc, 20 µg WbpM, 36 µg His_6-WreQ and 1 mM NADH. The lyophilized reaction mixture was treated with 1 M hydrochloric acid for 15 min at 100 °C. After reduction with $NaBD_4$ and acetylation, alditol acetate derivatives of sugars in the reaction were dissolved in dichloromethane and injected onto a HP-5MS (Agilent 19091S-433E) capillary column (0.25 mm inner diameter × 30 m, 0.25-µm film thickness) with helium as carrier gas at a flow rate of 1.0 ml/min. The GC program was started with an oven temperature of 50 °C, followed by an increase of 20 °C min^{-1} to 150 °C, then at 2 °C min^{-1} to 200 °C, then at 5 °C min^{-1} to 260 °C, and a hold at 260 °C for 10 min.

N. Purification of nucleotide sugars by thin layer chromatography

To purify the non-radioactive WbpM-reaction product for NMR analysis, the reaction was incubated at 30 °C for 1 hour with a total volume of 500 µl consisting of buffer D (20 mM sodium phosphate, pH=7.0), 1 mM UDP-GlcNAc and 120 µg WbpM. To purify the final product

of the WbpM-WreQ reaction, 450 µg His_6-WreQ and 1 mM NADH was added to the WbpM reaction and further incubated for 4 hours. Reactions were stopped and treated as described above, and the dried residues after lyophilization were dissolved in a 50 µl volume in water. Of this solution, 24 µl (0.24 µmol) was spotted on a TLC plate (ten spots, 2 µl each) and developed in Solvent A as described above. One loaded lane of the TLC plate was excised from both sides of the TLC plate, sprayed with thymol reagent (5 mg/ml thymol in ethanol/concentrated sulfuric = 19:1 (v/v)), and incubated in a 110 °C oven until stained spots were observed, in order to locate the position of the sugar nucleotide to be purified. At this position on the unstained part of the same TLC plate, the silica coating was scraped from the aluminum backing and the sugar nucleotide candidates were eluted from the silica with Solvent A. The eluates were lyophilized and re-dissolved in 20 µl water. The result of purification was analyzed by TLC with thymol staining and the concentration was measured by UV absorbance at 260 nm.

O. NMR spectroscopy

NMR experiments were performed at the static magnetic field strength of 14.1 Tesla (proton Larmor frequency of 600 MHz) on the Agilent VNMRS instrument equipped with the ColdProbe. Standard TOSCY, gCOSY, and NOESY two-dimensional pulse sequences were used for assignments; the pulse sequences were obtained from VNMRJ 2.0 software suite. Raw NMR data were processed using NMRPipe (104) and analyzed in Sparky (105). The NMR assignment experiments were performed at 30 °C to match conditions of the enzyme reaction directly monitored by NMR. In the assignment experiments 99% D_2O was used as a solvent with presaturation of the residual water signal. The chemical shifts measured at 30 °C were referenced to the water resonance because its temperature dependence is well known: water chemical shift was calculated to be at 4.700 ppm at 30 °C (106,107). As an additional reference, we recorded the

chemical shift of the acetone methyl group in water at the same temperature, which was 2.201±0.008 ppm.

NMR samples were prepared by dissolution of lyophilized powders in 99.8% D_2O (Cambridge Isotope Laboratories). The final concentration of UDP-QuiNAc in the Shigemi tube sample was 0.23 mM. The one-dimensional proton spectrum was acquired with water presaturation, 16,384 points and 128 transients (8 minutes of total acquisition time). All two-dimensional experiments were acquired with 16,384 points in direct and 512 points in indirect dimensions. The 2D ^{1}H-^{1}H gCOSY was acquired in the real mode with 32 transients (5 hours). The 2D ^{1}H-^{1}H TOCSY was acquired with 16 transients (11 hours). The 2D ^{1}H-^{1}H NOESY was acquired with 16 transients in duplicate (total of 24 hours). The final concentration of the 4-keto intermediate was significantly lower than that of UDP-QuiNAc, which was partially compensated by increased acquisition time.

P. *In vitro* WreU, WreU-WreQ coupled, and WreU-WreQ-WreG coupled enzyme assay

Synthesis of lipid substrate Und-P— To prepare ^{32}P-labeled Und-P for [^{32}P]-WreU enzyme assay, 3 µl dimethyl sulfoxide (DMSO) and 10 µl 10% Triton X-100 were added to a tube containing 13 nmol of dried undecaprenol (American Radiolabeled Chemicals). The tube was vortexed ensure solubilization of the lipid. To the same tube, 5 µM [γ-^{32}P]-ATP (2000 mCi/mmol) (PerkinElmer), 1 µl of purified DGK, 10 µl 0.5 mM Tris buffer, pH 8.0, 40 mM $MgCl_2$, and dH_2O were added to a total volume of 100 µl.

To prepare Und-P for [^{3}H]-WreU enzyme assay, 0.1 mM unlabeled ATP was used instead, other components of the reaction were the same. The DGK reactions were incubated at 30 °C and allowed to proceed for 1 hour before adding components for other enzymes assays.

WreU glycosyltransferase assay— 1 µl His_6-WreU-containing membrane fraction was added as crude WreU enzyme to the DGK reaction. In the [^{32}P]-WreU assay, nucleotide sugar

substrates tested were: UDP-GlcNAc (Sigma), WbpM reaction product (contained pure UDP-ADHexu), WbpM-WreQ reaction product (contained a mixture of UDP-ADHexu and UDP-QuiNAc), purified UDP-ADHexu, and purified UDP-QuiNAc. The concentration of each nucleotide sugar substrate was 0.05 mM.

In the [^{3}H]-WreU assay, WreU crude enzyme and ^{3}H-labeled nucleotide sugars were added to unlabeled DGK reaction. Nucleotide sugar substrates tested were UDP-[^{3}H]GlcNAc (Sigma), [^{3}H]-WbpM reaction product (contained pure UDP-[^{3}H]ADHexu), [^{3}H]-WbpM-WreQ reaction product (contained a mixture of UDP-[^{3}H]ADHexu and UDP-[^{3}H]QuiNAc). The concentration of each nucleotide sugar substrate was 0.02 mM.

The WreU reactions were incubated at 30 °C for 1 hr, then quenched into 500 µl of Solvent I (chloroform-methanol/3:2) and extracted with 400 µl PSUP (chloroform-methanol-1M $MgCl_2$-water/18:294:293:1) (108). The organic layers were dried with lyophilization and re-dissolved in 20 µl Solvent I. 1 µl of each sample was spotted on an aluminum-backed precoated Silica gel 60 plate (EMD Chemicals) and developed in TLC Solvent A. Dried TLC plates were exposed to films or photostimulable phosphor (PSP) plates and viewed by autoradiogram. For the [^{3}H]-WreU assay, aqueous layers of each reaction were also lyophilized, re-dissolved and the entire volume analyzed by TLC to view nucleotide sugars that did not participate in the reaction.

WreU-WreQ coupled assay—To the WreU reactions, 0.1 mM NADH and 1 µl (~ 10 µg) WreQ enzyme were added. The WreQ reactions were allowed to proceed for 1 hour at 30 °C, then quenched and prepared for analysis as described above for WreU enzyme assay.

To understand the rate of the WreQ reaction with the lipidated substrate, 0.1 mM NADH and 1 µl serially-diluted WreQ enzyme (10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1} and undiluted) was added to [^{32}P]-WreU reactions. The WreQ reactions were allowed for only 1 minute. To one WreU reaction, Solvent I was added before adding WreQ enzyme as a 0 minute control, for showing the effectiveness of the method of quenching the reactions. The organic phases of reaction samples

were analyzed by TLC expose. Radioactive (^{32}P) spots were quantified by phosphorimager and used for estimating reaction kinetics.

WreU-WreQ-WreG coupled reaction—In reactions that aimed to test the glycosyltransferase activity of WreG, 10 µl crude or purified WreG enzyme and 0.1 mM GDP-mannose were added to [^{32}P]-WreU reactions together with (or without) 10 µg WreQ and 0.1 mM NADH. Reactions were allowed for 1 hour after adding WreG and then quenched and prepared for TLC analysis. For analyzing the WreU-WreQ-WreG result, two TLC solvents were used: Solvent A and Solvent B (choloform/methanol/water, 65:25:4).

Q. Suppression test of *R. etli wreQ* mutant CE166

To test whether multicopy dosage of the *wreU* and/or the *wreG* gene could suppress the *wreQ*::Tn*5* mutant CE166, pLS22 plasmid carrying the *His$_6$-wreU* gene and pTL63 plasmid carrying the *wreG-His$_6$* gene were individually transferred into CE166 by a triparental mating procedure described previously (21). The LPS from the resulting strains were prepared and analyzed by SDS-PAGE with periodate-silver stain.

REFERENCES

1. Raetz, C. R. (1990) Biochemistry of endotoxins. *Annu Rev Biochem* **59**, 129-170

2. Raetz, C. R., and Whitfield, C. (2002) Lipopolysaccharide endotoxins. *Annu Rev Biochem* **71**, 635-700

3. Bhat, U. R., Forsberg, L. S., and Carlson, R. W. (1994) Structure of lipid A component of *Rhizobium leguminosarum* bv. *phaseoli* lipopolysaccharide. Unique nonphosphorylated lipid A containing 2-amino-2-deoxygluconate, galacturonate, and glucosamine. *J Biol Chem* **269**, 14402-14410

4. Forsberg, L. S., Bhat, U. R., and Carlson, R. W. (2000) Structural characterization of the O-antigenic polysaccharide of the lipopolysaccharide from *Rhizobium etli* strain CE3. A unique O-acetylated glycan of discrete size, containing 3-O-methyl-6-deoxy-L -talose and 2,3,4-tri-O-methyl-L-fucose. *J Biol Chem* **275**, 18851-18863

5. Forsberg, L. S., and Carlson, R. W. (1998) The structures of the lipopolysaccharides from *Rhizobium etli* strains CE358 and CE359. The complete structure of the core region of *R. etli* lipopolysaccharides. *J Biol Chem* **273**, 2747-2757

6. Cava, J. R., Elias, P. M., Turowski, D. A., and Noel, K. D. (1989) *Rhizobium leguminosarum* CFN42 genetic regions encoding lipopolysaccharide structures essential for complete nodule development on bean plants. *J Bacteriol* **171**, 8-15

7. Burns, S. M., and Hull, S. I. (1998) Comparison of loss of serum resistance by defined lipopolysaccharide mutants and an acapsular mutant of uropathogenic *Escherichia coli* O75:K5. *Infect Immun* **66**, 4244-4253

8. Joiner, K. A. (1988) Complement evasion by bacteria and parasites. *Annu Rev Microbiol* **42**, 201-230

9. Eisenschenk, L., Diebold, R., Perez-Lesher, J., Peterson, A. C., Kent Peters, N., and Noel, K. D. (1994) Inhibition of *Rhizobium etli* Polysaccharide Mutants by *Phaseolus vulgaris* Root Compounds. *Appl Environ Microbiol* **60**, 3315-3322

10. Ormeno-Orrillo, E., Rosenblueth, M., Luyten, E., Vanderleyden, J., and Martinez-Romero, E. (2008) Mutations in lipopolysaccharide biosynthetic genes impair maize rhizosphere and root colonization of *Rhizobium tropici* CIAT899. *Environ Microbiol* **10**, 1271-1284

11. Reuhs, B. L., Relic, B., Forsberg, L. S., Marie, C., Ojanen-Reuhs, T., Stephens, S. B., Wong, C. H., Jabbouri, S., and Broughton, W. J. (2005) Structural characterization of a flavonoid-inducible *Pseudomonas aeruginosa* A-band-like O antigen of *Rhizobium* sp. strain NGR234, required for the formation of nitrogen-fixing nodules. *J Bacteriol* **187**, 6479-6487

12. Ojeda, K. J., Simonds, L., and Noel, K. D. (2013) Roles of predicted glycosyltransferases in the biosynthesis of the *Rhizobium etli* CE3 O antigen. *J Bacteriol* **195**, 1949-1958

13. Bouhss, A., Trunkfield, A. E., Bugg, T. D., and Mengin-Lecreulx, D. (2008) The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiol Rev* **32**, 208-233

14. Whitfield, C., and Paiment, A. (2003) Biosynthesis and assembly of Group 1 capsular polysaccharides in *Escherichia coli* and related extracellular polysaccharides in other bacteria. *Carbohydr Res* **338**, 2491-2502

15. Gonzalez, V., Santamaria, R. I., Bustos, P., Hernandez-Gonzalez, I., Medrano-Soto, A., Moreno-Hagelsieb, G., Janga, S. C., Ramirez, M. A., Jimenez-Jacinto, V., Collado-Vides, J., and Davila, G. (2006) The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. *Proc Natl Acad Sci U S A* **103**, 3834-3839

16. Lerouge, I., Laeremans, T., Verreth, C., Vanderleyden, J., Van Soom, C., Tobin, A., and Carlson, R. W. (2001) Identification of an ATP-binding cassette transporter for export of the O-antigen across the inner membrane in *Rhizobium etli* based on the genetic, functional, and structural analysis of an lps mutant deficient in O-antigen. *J Biol Chem* **276**, 17190-17198

17. Samuel, G., and Reeves, P. (2003) Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. *Carbohydr Res* **338**, 2503-2519

18. Noel, K. D. (1992) Rhizobial polysaccharides required in symbioses with legumes. in *Molecular signals in plantmicrobe communications* (Verma, D. P. S. ed.), CRC Press, Boca Raton, Fla. pp 341-357

19. Forsberg, L. S., Noel, K. D., Box, J., and Carlson, R. W. (2003) Genetic locus and structural characterization of the biochemical defect in the O-antigenic polysaccharide of the symbiotically deficient *Rhizobium etli* mutant, CE166. Replacement of *N*-acetylquinovosamine with its hexosyl-4-ulose precursor. *J Biol Chem* **278**, 51347-51359

20. Garcia-de los Santos, A., and Brom, S. (1997) Characterization of two plasmid-borne *lpsβ* loci of *Rhizobium etli* required for lipopolysaccharide synthesis and for optimal interaction with plants. *Mol Plant Microbe Interact* **10**, 891-902

21. Ojeda, K. J. (2009) *Biosynthesis of the Rhizobium etli CE3 O antigen*Dissertation, Marquette University

22. Reeves, P. R., Hobbs, M., Valvano, M. A., Skurnik, M., Whitfield, C., Coplin, D., Kido, N., Klena, J., Maskell, D., Raetz, C. R., and Rick, P. D. (1996) Bacterial polysaccharide synthesis and gene nomenclature. *Trends Microbiol* **4**, 495-503

23. Alexander, D. C., and Valvano, M. A. (1994) Role of the rfe gene in the biosynthesis of the *Escherichia coli* O7-specific lipopolysaccharide and other O-specific polysaccharides containing *N*-acetylglucosamine. *J Bacteriol* **176**, 7079-7084

24. Wang, L., Liu, D., and Reeves, P. R. (1996) C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. *J Bacteriol* **178**, 2598-2604

25. Lehrer, J., Vigeant, K. A., Tatar, L. D., and Valvano, M. A. (2007) Functional characterization and membrane topology of *Escherichia coli* WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial common antigen and O-antigen lipopolysaccharide. *J Bacteriol* **189**, 2618-2628

26. Al-Dabbagh, B., Mengin-Lecreulx, D., and Bouhss, A. (2008) Purification and characterization of the bacterial UDP-GlcNAc:undecaprenyl-phosphate GlcNAc-1-phosphate transferase WecA. *J Bacteriol* **190**, 7141-7146

27. Saldias, M. S., Patel, K., Marolda, C. L., Bittner, M., Contreras, I., and Valvano, M. A. (2008) Distinct functional domains of the *Salmonella enterica* WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit. *Microbiology* **154**, 440-453

28. Patel, K. B., Furlong, S. E., and Valvano, M. A. (2010) Functional analysis of the C-terminal domain of the WbaP protein that mediates initiation of O antigen synthesis in *Salmonella enterica*. *Glycobiology* **20**, 1389-1401

29. Patel, K. B., Ciepichal, E., Swiezewska, E., and Valvano, M. A. (2012) The C-terminal domain of the *Salmonella enterica* WbaP (UDP-galactose:Und-P galactose-1-phosphate transferase) is sufficient for catalytic activity and specificity for undecaprenyl monophosphate. *Glycobiology* **22**, 116-122

30. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539

31. Carlson, R. W., Reuhs, B., Chen, T. B., Bhat, U. R., and Noel, K. D. (1995) Lipopolysaccharide core structures in *Rhizobium etli* and mutants deficient in O-antigen. *J Biol Chem* **270**, 11783-11788

32. Kadrmas, J. L., Allaway, D., Studholme, R. E., Sullivan, J. T., Ronson, C. W., Poole, P. S., and Raetz, C. R. (1998) Cloning and overexpression of glycosyltransferases that generate the lipopolysaccharide core of *Rhizobium leguminosarum*. *J Biol Chem* **273**, 26432-26440

33. Kanjilal-Kolar, S., Basu, S. S., Kanipes, M. I., Guan, Z., Garrett, T. A., and Raetz, C. R. (2006) Expression cloning of three *Rhizobium leguminosarum* lipopolysaccharide core galacturonosyltransferases. *J Biol Chem* **281**, 12865-12878

34. Allaway, D., Jeyaretnam, B., Carlson, R. W., and Poole, P. S. (1996) Genetic and chemical characterization of a mutant that disrupts synthesis of the lipopolysaccharide core tetrasaccharide in *Rhizobium leguminosarum*. *J Bacteriol* **178**, 6403-6406

35. Noel, K. D., Forsberg, L. S., and Carlson, R. W. (2000) Varying the abundance of O antigen in *Rhizobium etli* and its effect on symbiosis with *Phaseolus vulgaris*. *J Bacteriol* **182**, 5317-5324

36. Hao, Y., and Lam, J. (2011) Pathways for the Biosynthesis of NDP Sugars. in *Bacterial Lipopolysaccharides* (Knirel, Y. A., and Valvano, M. A. eds.), Springer Vienna. pp 195-235

37. Creuzenet, C., and Lam, J. S. (2001) Topological and functional characterization of WbpM, an inner membrane UDP-GlcNAc C6 dehydratase essential for lipopolysaccharide biosynthesis in *Pseudomonas aeruginosa*. *Mol Microbiol* **41**, 1295-1310

38. Olivier, N. B., Chen, M. M., Behr, J. R., and Imperiali, B. (2006) In vitro biosynthesis of UDP-*N*,*N'*-diacetylbacillosamine by enzymes of the *Campylobacter jejuni* general protein glycosylation system. *Biochemistry* **45**, 13659-13669

39. Schoenhofen, I. C., McNally, D. J., Vinogradov, E., Whitfield, D., Young, N. M., Dick, S., Wakarchuk, W. W., Brisson, J. R., and Logan, S. M. (2006) Functional characterization of dehydratase/aminotransferase pairs from *Helicobacter* and *Campylobacter*: enzymes distinguishing the pseudaminic acid and bacillosamine biosynthetic pathways. *J Biol Chem* **281**, 723-732

40. Pinta, E., Duda, K. A., Hanuszkiewicz, A., Kaczynski, Z., Lindner, B., Miller, W. L., Hyytiainen, H., Vogel, C., Borowski, S., Kasperkiewicz, K., Lam, J. S., Radziejewska-Lebrecht, J., Skurnik, M., and Holst, O. (2009) Identification and role of a 6-deoxy-4-keto-hexosamine in the lipopolysaccharide outer core of *Yersinia enterocolitica* serotype O:3. *Chemistry* **15**, 9747-9754

41. Creuzenet, C., Urbanic, R. V., and Lam, J. S. (2002) Structure-function studies of two novel UDP-GlcNAc C6 dehydratases/C4 reductases. Variation from the SYK dogma. *J Biol Chem* **277**, 26769-26778

42. Kavanagh, K. L., Jornvall, H., Persson, B., and Oppermann, U. (2008) Medium- and short-chain dehydrogenase/reductase gene and protein families : the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. *Cell Mol Life Sci* **65**, 3895-3906

43. Belanger, M., Burrows, L. L., and Lam, J. S. (1999) Functional analysis of genes responsible for the synthesis of the B-band O antigen of *Pseudomonas aeruginosa* serotype O6 lipopolysaccharide. *Microbiology* **145 (Pt 12)**, 3505-3521

44. Greenfield, L. K., Richards, M. R., Li, J., Wakarchuk, W. W., Lowary, T. L., and Whitfield, C. (2012) Biosynthesis of the polymannose lipopolysaccharide O-antigens from *Escherichia coli* serotypes O8 and O9a requires a unique combination of single- and multiple-active site mannosyltransferases. *J Biol Chem* **287**, 35078-35091

45. Knirel, Y. A., Bystrova, O. V., Kocharova, N. A., Zahringer, U., and Pier, G. B. (2006) Conserved and variable structural features in the lipopolysaccharide of *Pseudomonas aeruginosa*. *J Endotoxin Res* **12**, 324-336

46. Bystrova, O. V., Shashkov, A. S., Kocharova, N. A., Knirel, Y. A., Lindner, B., Zahringer, U., and Pier, G. B. (2002) Structural studies on the core and the O-polysaccharide repeating unit of *Pseudomonas aeruginosa* immunotype 1 lipopolysaccharide. *Eur J Biochem* **269**, 2194-2203

47. Creuzenet, C., Schur, M. J., Li, J., Wakarchuk, W. W., and Lam, J. S. (2000) FlaA1, a new bifunctional UDP-GlcNAc C6 Dehydratase/ C4 reductase from *Helicobacter pylori*. *J Biol Chem* **275**, 34873-34880

48. Merrick, M., Filser, M., Kennedy, C., and Dixon, R. (1978) Polarity of mutations induced by insertion of transposons Tn5, Tn7 and Tn10 into the nif gene cluster of *Klebsiella pneumoniae*. *Mol Gen Genet* **165**, 103-111

49. Berg, D. E., Weiss, A., and Crossland, L. (1980) Polarity of Tn5 insertion mutations in *Escherichia coli*. *J Bacteriol* **142**, 439-446

50. Ojeda, K. J., Box, J. M., and Noel, K. D. (2010) Genetic basis for *Rhizobium etli* CE3 O-antigen O-methylated residues that vary according to growth conditions. *J Bacteriol* **192**, 679-690

51. Dombrecht, B., Vanderleyden, J., and Michiels, J. (2001) Stable RK2-derived cloning vectors for the analysis of gene expression and gene function in gram-negative bacteria. *Mol Plant Microbe Interact* **14**, 426-430

52. King, J. D., Poon, K. K., Webb, N. A., Anderson, E. M., McNally, D. J., Brisson, J. R., Messner, P., Garavito, R. M., and Lam, J. S. (2009) The structural basis for catalytic function of GMD and RMD, two closely related enzymes from the GDP-D-rhamnose biosynthesis pathway. *FEBS J* **276**, 2686-2700

53. Claridge, T. D. W. (2009) *High-Resolution NMR Techniques in Organic Chemistry*, Elsevier, Amsterdam

54. Smith, E. J. (1964) The isolation and characterization of 2-amino-2:6-dideoxy d-glucose (D-quinovosamine) from a bacterial polysaccharide. *Biochemical and Biophysical Research Communications* **15**, 593-597

55. Luderitz, O., Gmeiner, J., Kickhofen, B., Mayer, H., Westphal, O., and Wheat, R. W. (1968) Identification of D-mannosamine and quinovosamine in *Salmonella* and related bacteria. *J Bacteriol* **95**, 490-494

56. Daniel, A., Raff, R. A., and Wheat, R. W. (1972) Identification of products of the uridinediphospho-*N*-acetyl-D-glucosamine oxidoreductase system from *Citrobacter freundii* ATCC 10053. *J Bacteriol* **110**, 110-116

57. Bonin, C. P., Potter, I., Vanzin, G. F., and Reiter, W. D. (1997) The MUR1 gene of *Arabidopsis thaliana* encodes an isoform of GDP-D-mannose-4,6-dehydratase, catalyzing the first step in the de novo synthesis of GDP-L-fucose. *Proc Natl Acad Sci U S A* **94**, 2085-2090

58. Ohyama, C., Smith, P. L., Angata, K., Fukuda, M. N., Lowe, J. B., and Fukuda, M. (1998) Molecular cloning and expression of GDP-D-mannose-4,6-dehydratase, a key enzyme for fucose metabolism defective in Lec13 cells. *J Biol Chem* **273**, 14582-14587

59. McNally, D. J., Schoenhofen, I. C., Mulrooney, E. F., Whitfield, D. M., Vinogradov, E., Lam, J. S., Logan, S. M., and Brisson, J. R. (2006) Identification of labile UDP-ketosugars in *Helicobacter pylori*, *Campylobacter jejuni* and *Pseudomonas aeruginosa*: key metabolites used to make glycan virulence factors. *Chembiochem* **7**, 1865-1868

60. Merino, S., Jimenez, N., Molero, R., Bouamama, L., Regue, M., and Tomas, J. M. (2011) A UDP-HexNAc:polyprenol-P GalNAc-1-P transferase (WecP) representing a new subgroup of the enzyme family. *J Bacteriol* **193**, 1943-1952

61. Burrows, L. L., Charter, D. F., and Lam, J. S. (1996) Molecular characterization of the *Pseudomonas aeruginosa* serotype O5 (PAO1) B-band lipopolysaccharide gene cluster. *Mol Microbiol* **22**, 481-495

62. Hartley, M. D., Larkin, A., and Imperiali, B. (2008) Chemoenzymatic synthesis of polyprenyl phosphates. *Bioorg Med Chem* **16**, 5149-5156

63. El Ghachi, M., Bouhss, A., Blanot, D., and Mengin-Lecreulx, D. (2004) The bacA gene of *Escherichia coli* encodes an undecaprenyl pyrophosphate phosphatase activity. *J Biol Chem* **279**, 30106-30113

64. Valvano, M. A. (2003) Export of O-specific lipopolysaccharide. *Front Biosci* **8**, s452-471

65. Glover, K. J., Weerapana, E., Chen, M. M., and Imperiali, B. (2006) Direct biochemical evidence for the utilization of UDP-bacillosamine by PglC, an essential glycosyl-1-phosphate transferase in the *Campylobacter jejuni* N-linked glycosylation pathway. *Biochemistry* **45**, 5343-5350

66. Kanjilal-Kolar, S., and Raetz, C. R. (2006) Dodecaprenyl phosphate-galacturonic acid as a donor substrate for lipopolysaccharide core glycosylation in *Rhizobium leguminosarum*. *J Biol Chem* **281**, 12879-12887

67. Forsberg, L. S., and Carlson, R. W. (2008) Structural characterization of the primary O-antigenic polysaccharide of the *Rhizobium leguminosarum* 3841 lipopolysaccharide and identification of a new 3-acetimidoylamino-3-deoxyhexuronic acid glycosyl component: a unique O-methylated glycan of uniform size, containing 6-deoxy-3-O-methyl-D-talose, *N*-acetylquinovosamine, and rhizoaminuronic acid (3-acetimidoylamino-3-deoxy-D-gluco-hexuronic acid). *J Biol Chem* **283**, 16037-16050

68. Kneidinger, B., Larocque, S., Brisson, J. R., Cadotte, N., and Lam, J. S. (2003) Biosynthesis of 2-acetamido-2,6-dideoxy-L-hexoses in bacteria follows a pattern distinct from those of the pathways of 6-deoxy-L-hexoses. *Biochem J* **371**, 989-995

69. Muszynski, A., Laus, M., Kijne, J. W., and Carlson, R. W. (2010) Structures of the lipopolysaccharides from *Rhizobium leguminosarum* RBL5523 and its UDP-glucose dehydrogenase mutant (exo5). *Glycobiology* **21**, 55-68

70. Wang, Z., Larocque, S., Vinogradov, E., Brisson, J. R., Dacanay, A., Greenwell, M., Brown, L. L., Li, J., and Altman, E. (2004) Structural studies of the capsular polysaccharide and lipopolysaccharide O-antigen of *Aeromonas salmonicida* strain 80204-1 produced under *in vitro* and *in vivo* growth conditions. *Eur J Biochem* **271**, 4507-4516

71. De Castro, C., Molinaro, A., Lanzetta, R., Holst, O., and Parrilli, M. (2005) The linkage between O-specific caryan and core region in the lipopolysaccharide of *Burkholderia caryophylli* is furnished by a primer monosaccharide. *Carbohydr Res* **340**, 1802-1807

72. Vinogradov, E. V., Shashkov, A. S., Knirel, Y. A., Kochetkov, N. K., Tochtamysheva, N. V., Averin, S. F., Goncharova, O. V., and Khlebnikov, V. S. (1991) Structure of the O-antigen of *Francisella tularensis* strain 15. *Carbohydr Res* **214**, 289-297

73. Wang, Q., Shi, X., Leymarie, N., Madico, G., Sharon, J., Costello, C. E., and Zaia, J. (2011) A typical preparation of *Francisella tularensis* O-antigen yields a mixture of three types of saccharides. *Biochemistry* **50**, 10941-10950

74. Knirel, Y. A., Moll, H., and Zahringer, U. (1996) Structural study of a highly O-acetylated core of *Legionella pneumophila* serogroup 1 lipopolysaccharide. *Carbohydr Res* **293**, 223-234

75. Kondakova, A. N., Kirsheva, N. A., Shashkov, A. S., Shaikhutdinova, R. Z., Arabtsky, N. P., Ivanov, S. A., Anisimov, A. P., and Knirel, Y. A. (2011) Low structural diversity of the O-polysaccharides of *Photorhabdus asymbiotica* subspp. *asymbiotica* and *australis* and their similarity to the O-polysaccharides of taxonomically remote bacteria including *Francisella tularensis*. *Carbohydr Res* **346**, 1951-1955

76. Muldoon, J., Shashkov, A. S., Senchenkova, S. N., Tomshich, S. V., Komandrova, N. A., Romanenko, L. A., Knirel, Y. A., and Savage, A. V. (2001) Structure of an acidic polysaccharide from a marine bacterium *Pseudoalteromonas distincta* KMM 638 containing 5-acetamido-3,5,7,9-tetradeoxy-7-formamido-L-glycero-L-manno-nonulosonic acid. *Carbohydr Res* **330**, 231-239

77. Lam, J. S., Taylor, V. L., Islam, S. T., Hao, Y., and Kocincova, D. (2011) Genetic and Functional Diversity of *Pseudomonas aeruginosa* Lipopolysaccharide. *Front Microbiol* **2**, 118

78. Molinaro, A., Bedini, E., Ferrara, R., Lanzetta, R., Parrilli, M., Evidente, A., Lo Cantore, P., and Iacobellis, N. S. (2003) Structural determination of the O-specific chain of the lipopolysaccharide from the mushrooms pathogenic bacterium Pseudomonas tolaasii. *Carbohydr Res* **338**, 1251-1257

79. Vinogradov, E. V., Holst, O., Thomas-Oates, J. E., Broady, K. W., and Brade, H. (1992) The structure of the O-antigenic polysaccharide from lipopolysaccharide of *Vibrio cholerae* strain H11 (non-O1). *Eur J Biochem* **210**, 491-498

80. Knirel, Y. A., Paredes, L., Jansson, P. E., Weintraub, A., Widmalm, G., and Albert, M. J. (1995) Structure of the capsular polysaccharide of *Vibrio cholerae* O139 synonym Bengal containing D-galactose 4,6-cyclophosphate. *Eur J Biochem* **232**, 391-396

81. Perepelov, A. V., Kocharova, N. A., Knirel, Y. A., Jansson, P. E., and Weintraub, A. (2011) Structure of the O-polysaccharide of *Vibrio cholerae* O43 containing a new monosaccharide derivative, 4-(*N*-acetyl-L-allothreonyl)amino-4,6-dideoxy-D-glucose. *Carbohydr Res* **346**, 430-433

82. Reddy, G. P., Hayat, U., Abeygunawardana, C., Fox, C., Wright, A. C., Maneval, D. R., Jr., Bush, C. A., and Morris, J. G., Jr. (1992) Purification and determination of the structure of capsular polysaccharide of *Vibrio vulnificus* M06-24. *J Bacteriol* **174**, 2620-2630

83. Reddy, G. P., Hayat, U., Xu, Q., Reddy, K. V., Wang, Y., Chiu, K. W., Morris, J. G., Jr., and Bush, C. A. (1998) Structure determination of the capsular polysaccharide from *Vibrio vulnificus* strain 6353. *Eur J Biochem* **255**, 279-288

84. Dmitriev, B. A., Kocharova, N. A., Knirel, Y. A., Shashkov, A. S., Kochetkov, N. K., Stanislavsky, E. S., and Mashilova, G. M. (1982) Somatic Antigens of *Pseudomonas aeruginosa*. The structure of the polysaccharide chain of Ps.aeruginosa O:6 (Lanyi) lipopolysaccharide. *Eur J Biochem* **125**, 229-237

85. White, D. (2007) *The physiology and biochemistry of prokaryates* 3rd ed., NY: Oxford University Press Inc.

86. Breitling, J., and Aebi, M. (2013) N-linked protein glycosylation in the endoplasmic reticulum. *Cold Spring Harb Perspect Biol* **5**, a013359

87. Mostafavi, A. Z., and Troutman, J. M. (2013) Biosynthetic assembly of the *Bacteroides fragilis* capsular polysaccharide A precursor bactoprenyl diphosphate-linked acetamido-4-amino-6-deoxygalactopyranose. *Biochemistry* **52**, 1939-1949

88. Sadovskaya, I., Brisson, J. R., Khieu, N. H., Mutharia, L. M., and Altman, E. (1998) Structural characterization of the lipopolysaccharide O-antigen and capsular polysaccharide of *Vibrio ordalii* serotype O:2. *Eur J Biochem* **253**, 319-327

89. MacLean, L. L., Perry, M. B., Crump, E. M., and Kay, W. W. (2003) Structural characterization of the lipopolysaccharide O-polysaccharide antigen produced by *Flavobacterium columnare* ATCC 43622. *Eur J Biochem* **270**, 3440-3446

90. Shashkov, A. S., Senchenkova, S. N., Chizhov, A. O., Knirel, Y. A., Esteve, C., Alcaide, E., Merino, S., and Tomas, J. M. (2009) Structure of a polysaccharide from the lipopolysaccharides of *Vibrio vulnificus* strains CECT 5198 and S3-I2-36, which is remarkably similar to the O-polysaccharide of *Pseudoalteromonas rubra* ATCC 29570. *Carbohydr Res* **344**, 2005-2009

91. Pinta, E., Duda, K. A., Hanuszkiewicz, A., Salminen, T. A., Bengoechea, J. A., Hyytiainen, H., Lindner, B., Radziejewska-Lebrecht, J., Holst, O., and Skurnik, M. (2010) Characterization of the six glycosyltransferases involved in the biosynthesis of *Yersinia enterocolitica* serotype O:3 lipopolysaccharide outer core. *J Biol Chem* **285**, 28333-28342

92. Anderson, M. S., Eveland, S. S., and Price, N. P. (2000) Conserved cytoplasmic motifs that distinguish sub-groups of the polyprenol phosphate:N-acetylhexosamine-1-phosphate transferase family. *FEMS Microbiol Lett* **191**, 169-175

93. Price, N. P., and Momany, F. A. (2005) Modeling bacterial UDP-HexNAc: polyprenol-P HexNAc-1-P transferases. *Glycobiology* **15**, 29R-42R

94. Noel, K. D., Sanchez, A., Fernandez, L., Leemans, J., and Cevallos, M. A. (1984) *Rhizobium phaseoli* symbiotic mutants with transposon Tn5 insertions. *J Bacteriol* **158**, 148-155

95. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J Mol Biol* **215**, 403-410

96. Hoang, T. T., Karkhoff-Schweizer, R. R., Kutchma, A. J., and Schweizer, H. P. (1998) A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. *Gene* **212**, 77-86

97. Schweizer, H. D. (1993) Small broad-host-range gentamycin resistance gene cassettes for site-specific insertion and deletion mutagenesis. *Biotechniques* **15**, 831-834

98. Alexeyev, M. F. (1995) Three kanamycin resistance gene cassettes with different polylinkers. *Biotechniques* **18**, 52, 54, 56

99. Glazebrook, J., and Walker, G. C. (1991) Genetic techniques in *Rhizobium meliloti*. *Methods Enzymol* **204**, 398-418

100. Finan, T. M., Kunkel, B., De Vos, G. F., and Signer, E. R. (1986) Second symbiotic megaplasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. *J Bacteriol* **167**, 66-72

101. Blatny, J. M., Brautaset, T., Winther-Larsen, H. C., Haugan, K., and Valla, S. (1997) Construction and use of a versatile set of broad-host-range cloning and expression vectors based on the RK2 replicon. *Appl Environ Microbiol* **63**, 370-379

102. Bouhenni, R. A., Vora, G. J., Biffinger, J. C., Shirodkar, S., Brockman, K., Ray, R., Wu, P., Johnson, B. J., Biddle, E. M., Marshall, M. J., Fitzgerald, L. A., Little, B. J., Fredrickson, J. K., Beliaev, A. S., Ringeisen, B. R. and Saffarini, D. A. (2010) The Role of *Shewanella oneidensis* MR-1 Outer Surface Structures in Extracellular Electron Transfer. *Electroanalysis* **22**, 856-864

103. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-254

104. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293

105. Goddard, T. D., and Kneller, D. G. SPARKY 3; http://www.cgl.ucsf.edu/home/sparky/. University of California, San Francisco

106. Hartel, A. J., Lankhorst, P. P., and Altona, C. (1982) Thermodynamics of stacking and of self-association of the dinucleoside monophosphate m2(6)A-U from proton NMR chemical shifts: differential concentration temperature profile method. *Eur J Biochem* **129**, 343-357

107. Orbons, L. P., van der Marel, G. A., van Boom, J. H., and Altona, C. (1987) An NMR study of the polymorphous behavior of the mismatched DNA octamer d(m5C-G-m5C-G-T-G-m5C-G) in solution. The B, Z, and hairpin forms. *J Biomol Struct Dyn* **4**, 939-963

108. Patel, K. B., and Valvano, M. A. (2013) In vitro UDP-sugar:undecaprenyl-phosphate sugar-1-phosphate transferase assay and product detection by thin layer chromatography. *Methods Mol Biol* **1022**, 173-183

APPENDICES

A. Mutants defective in the *wreU* and *wreQ* genes of *Rhizobium leguminosarum* 3841

Genomic studies identified genes encoding WreU, WreV and WreQ homologs in *R. leguminosarum* 3841 (*RL*3841). Their functions are expected to be the same as their *R. etli* CE3 orthologs, which are responsible for the coordinated QuiNAc (D-configuration) synthesis and O-antigen initiation. However, as discussed in Chapter Four (Genomics), structural study of the O antigen of *RL*3841 identified QuiNAc in the L-configuration, but no D-QuiNAc (67). To further investigate this discrepancy, the roles of *RL*3841 *wreU* ($wreU_{3841}$) and *wreQ* ($wreQ_{3841}$) gene were investigated by site-directed mutagenesis. *wreV* has two almost identical copies (99% identity) in the genome of 3841 and their mutagenesis was not pursued.

When analyzed on SDS-PAGE, *RL*3841 LPS was separated into three major bands (Fig. 42, lane 1), a high molecular weight LPS (smooth LPS or S-LPS), a low molecular weight LPS (rough LPS or R-LPS) and a medium molecular weight LPS (which we call it SR-LPS here). The S-LPS and R-LPS of *RL*3841 are comparable to the LPS I and LPS II of *R. etli* CE3, respectively (Fig. 42, lane 4). Although the composition of these *RL*3841 LPS species is not known, it can be proposed that the S-LPS is composed of all three domains of LPS, whereas the R-LPS contains no O antigen but only the lipid A and core. The SR-LPS may contain a truncated O-antigen with fewer repeating units.

In the $wreU_{3841}$ mutant strain LT11 ($wreU_{3841}$::Km), the O-antigen-containing S-LPS and SR-LPS were lost, and the amount of O-antigen-less R-LPS increases (Fig. 42, lane 2). This phenotype matches the *R. etli wreU* mutant in which LPS I is lost and the amount of LPS II increases. For complementation, the *R. etli* CE3 *wreU* gene was introduced into this mutant on a plasmid, and the LPS phenotype of this mutant was restored to wild type (Fig. 42, lane 3).

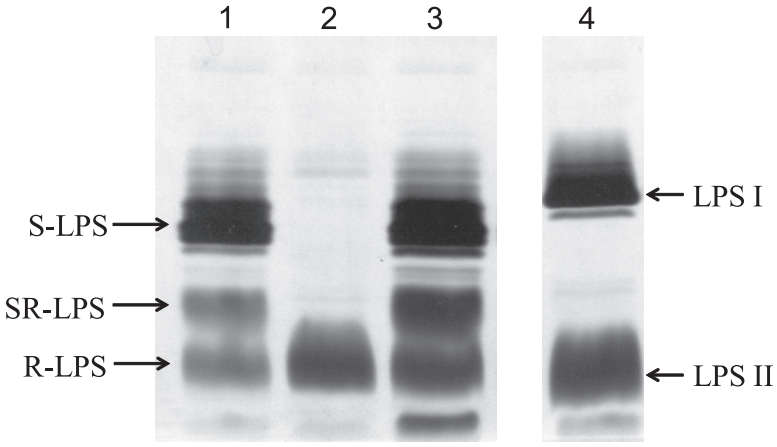


Figure 42. SDS-PAGE analysis of the result of *wreU*$_{3841}$ mutagenesis and complementation of the mutant with *R. etli* CE3 *wreU gene*. Lanes: lane 1, *R. leguminosarum* 3841; lane 2, LT11 (*wreU*$_{3841}$::Km); lane 3, LT11/pKT100 (pKT100 is the pJB21 plasmid carrying the *R. etli* CE3 *wreU* gene; lane 4, *R. etli* CE3. S-LPS and LPS I, LPS that contain lipid A, core and O-antigen; R-LPS and LPS II, LPS that contain lipid A and core, but no O-antigen; SR-LPS, LPS that contain lipid A, core, and a truncated O-antigen.

The LPS profile of the *wreQ*$_{3841}$ mutant strain LT9 (*wreQ*$_{3841}$::Km) (Fig. 43A, lane 2) is distinct from both the wild type and the *wreU*$_{3841}$ mutant (Fig. 43A, lane 1 and 3). The wild-type S-LPS band was lost in the *wreQ*$_{3841}$ mutant, accompanied with an increased staining intensity in the lower molecular weight region, overlapping with the SR-LPS. This observation suggests that the *wreQ*$_{3841}$ mutant may produce LPS with truncated O-antigens linked to the lipid A-core. Immunoblotting with antibody for *RL*3841 detected the S-LPS in the wild type (Fig. 43B, lane 1). The R-LPS in the *wreU*$_{3841}$ mutant was not detected, indicating that the epitope for detection requires structural features conferred by the O-antigen. At least some of the truncated LPSs of the *wreQ*$_{3841}$ mutant were detected by antibody, supporting that O-antigen is still made in this strain.

In conclusion, mutagenesis of *wreU*$_{3841}$ and *wreQ*$_{3841}$ both caused mutant LPS phenotypes in *RL*3841, which confirmed their involvement in O-antigen synthesis. WreU$_{3841}$ is mostly likely to be the initiating GTase in O-antigen synthesis, as the disruption of *wreU*$_{3841}$ abolished O-antigen synthesis. WreQ$_{3841}$ is expected to catalyze the reduction of ADHexu to QuiNAc on the Und-PP linkage, as determined for WreQ of *R. etli* CE3. O-antigens are still made in the *wreQ*$_{3841}$ mutant, but with a truncated length. It can be speculated that QuiNAc is replaced by ADHexu in

the truncated O-antigens. It is unclear how this replacement affects the O-antigen synthesis in *RL*3841.

The results of $wreU_{3841}$ and $wreQ_{3841}$ mutagenesis provide support from the genetic aspect that the QuiNAc in *RL*3841 O antigen may be actually in D-configuration. If it is in L-configuration, the absence of WreQ activity is not expected to affect O-antigen synthesis. Moreover, WreU is the initiating GTase for O-antigens with D-QuiNAc as the first sugar.

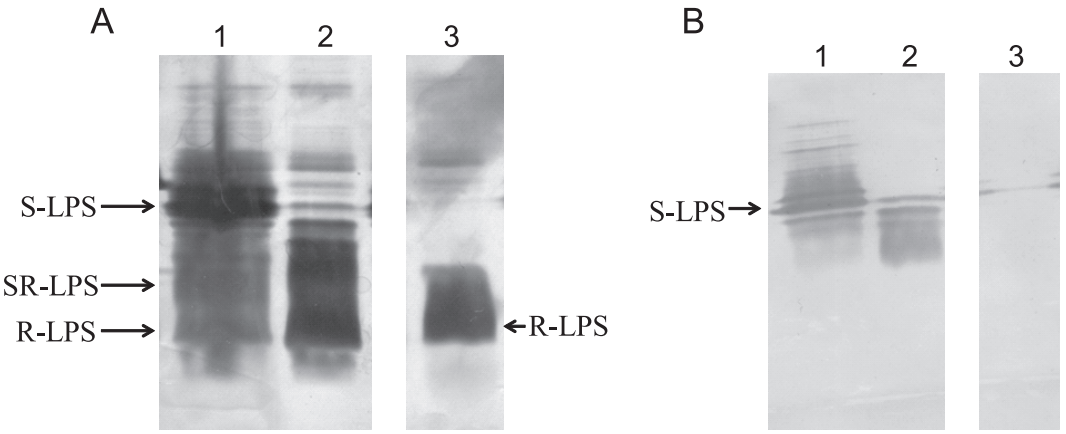


Figure 43. SDS-PAGE and immunoblot analysis of the LPS of *R. leguminosarum* 3841 mutants. (A) Non-blotted 18% SDS-PAGE gel with LPS visualized by silver staining. (B) Immunoblot of a gel similar to the one shown in (*A*) decorated with anti-3841 antisera. Lanes: lane 1, *R. leguminosarum* 3841; lane 2, LT9 ($wreQ_{3841}$::Km); lane 3, LT11 ($wreU_{3841}$::Km). SDS-PAGE and immunoblot were performed by Laurie Simonds.

B. Mutants defective in the O-antigen ligase genes of *R. etli* CE3

Until now, all ORFs within the three O-antigen genetic clusters in *R. etli* CE3 (shown in Fig. 6) have been mutated, with the exception of one gene hypothesized to encode an O-antigen ligase. This gene is located on plasmid B downstream of the *wreV* gene and is transcribed in an opposite direction. It had been temporarily given the name *orfL* in a previous study (21). When using the amino acid sequence of this gene product in a BLAST search against *R. etli* CE3, a homolog was found which is encoded by a chromosomal gene. The two genes are given new names in the present study: *orfL* is renamed oal_{pB} (oal, O-antigen ligase; $_{pB}$, plasmid B) and the

chromosome-located homolog is named *oal*$_{chr}$ ($_{chr}$, chromosome). Alignment of the two gene products is shown in Fig. 44 and the location of the two genes in the genome is shown in Fig. 45.

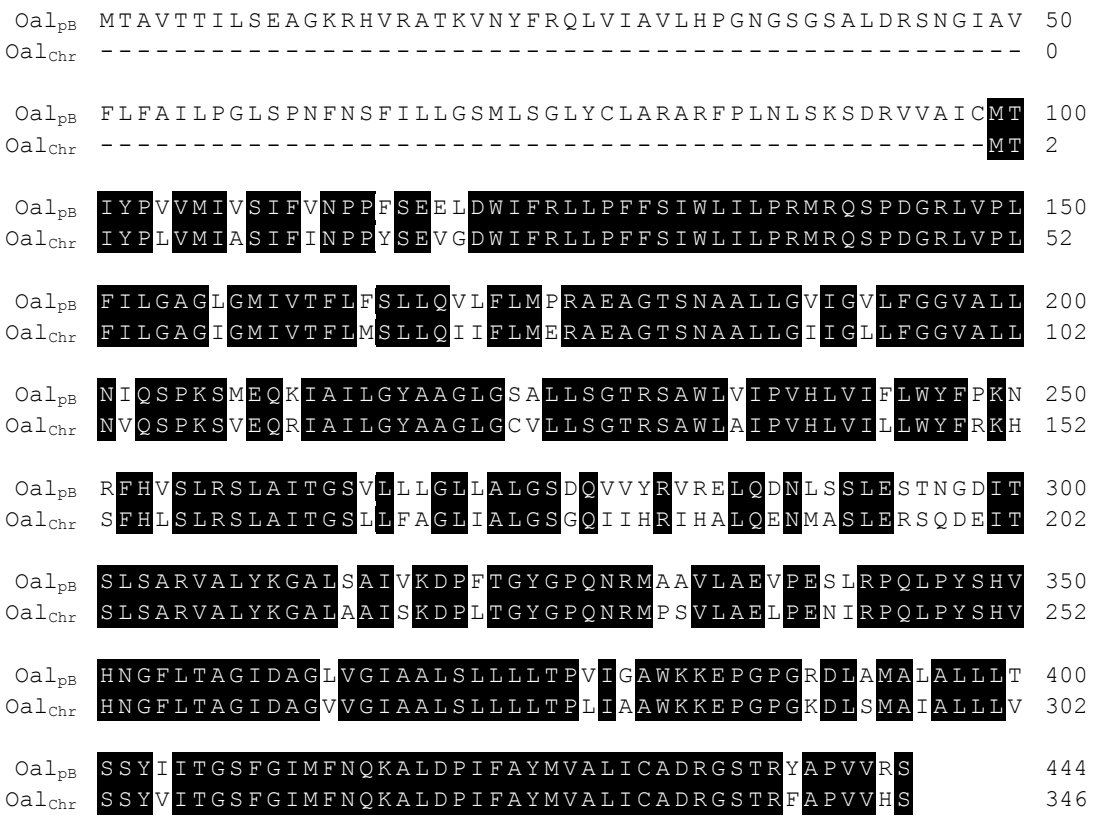


Figure 44. Clustal alignments of the two predicted O-antigen ligases encoded in the genome of *R. etli* CE3. Oal$_{pB}$ is encoded by the *oal*$_{pB}$ gene on plasmid B. Oal$_{chr}$ is encoded by the *oal*$_{chr}$ gene on the chromosome. Identical amino acids are highlighted.

***Rhizobium etli* CE3**

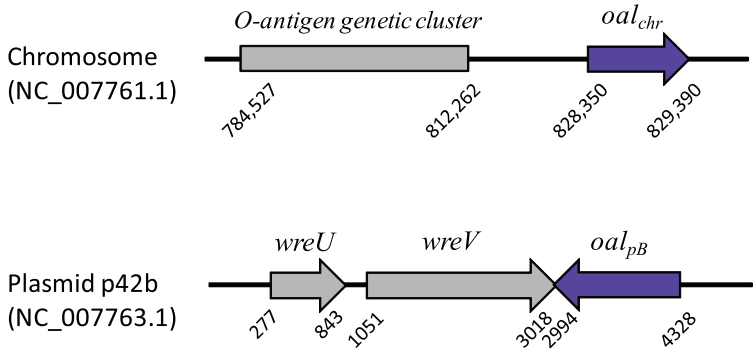


Figure 45. Schematic diagram showing the location of the two O-antigen ligase genes in the genome of *R. etli* CE3. The chromosomal copy of the *oal* gene is located near the 28 kb O-antigen genetic cluster, but is still separated from it by about 16 kb. The plasmid copy of *oal* gene is located immediately downstream of the *wreV* gene, and it is transcribed in an opposite direction.

O-antigen ligases catalyze the transfer of the O-antigen from an Und-PP linked O-antigen intermediate to the lipid A-core. This reaction can be considered as the last step of O-antigen synthesis. Because QuiNAc is the sugar whose linkage is changed in the ligase reaction for CE3 LPS, this step can be considered the end of the chemical journey of the QuiNAc moiety in the *R. etli* cell. Thereby the ligase reaction is logically connected to the rest of this dissertation.

The LPS of an O-antigen ligase mutant is expected to lack O-antigen completely. To investigate the role of the two putative O-antigen ligase genes in O-antigen synthesis in *R. etli* CE3, the mutagenesis strategy was to create a single mutant for each gene, as well as a double mutant in which both genes were mutated.

The result of mutagenesis is shown in Fig. 46. In the two single mutants, LPS I (O-antigen containing LPS) was present (Fig. 46, lane 2 and 3), although its amount was arguably less than in the wild type (Fig. 46, lane 1). It is more obvious that the amount of LPS II (lipid A-core only) increased in both single mutants. The double mutant is completely devoid of LPS I and its LPS II is more abundant (Fig. 46, lane 4).

The result of mutagenesis supported that the two genes encode O-antigen ligases for *R. etli* CE3 because when both genes are mutated, no LPS I (O-antigen containing LPS) is present.

The result also suggests that both genes can be expressed and are functional in *R. etli* CE3, because mutating one of them did not result in the loss of LPS I. However, the ratio of LPS II/LPS I increased substantially in both of the single mutants, indicating that both of the two O-antigen ligases are required for maintaining the normal LPS phenotype in the wild type. In all sequenced *R. etli* and *R. leguminosarum* strains, only *R. etli* CE3 contains two O-antigen ligase genes, whereas the other strains only contain the plasmid-encoded gene downstream of *wreV*. The chromosomal O-antigen ligase gene in *R. etli* CE3 may have resulted from a recent gene duplication and transposition event. The relationship between having two O-antigen ligase genes and the O-antigen synthesis in *R. etli* CE3 is unclear.

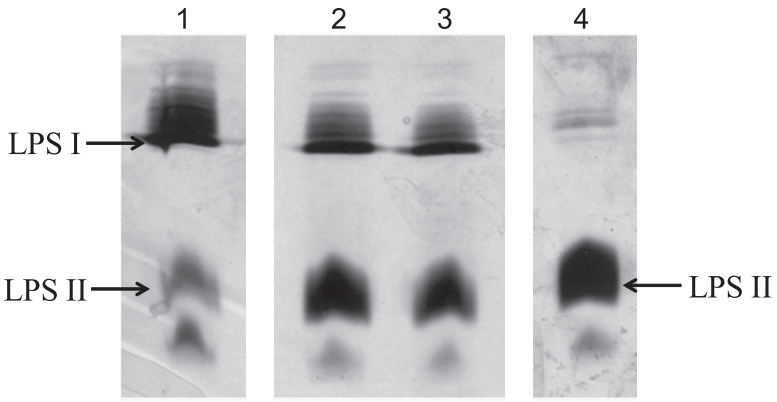


Figure 46. SDS-PAGE analysis of LPS of the *R. etli* O-antigen ligase single mutant and double mutant strains. Lanes: lane 1, CE3 (wildtype); lane 2, CE575 (oal_{pb}::Gm), plasmid B O-antigen ligase gene mutant; lane 3, CE577 (oal_{chr}::Gm), chromosomal O-antigen ligase gene mutant; lane 4, CE578 (oal_{pb}::Gm, oal_{chr}::Km), O-antigen ligase double mutant. This SDS-PAGE is performed by Laurie Simonds.

Methods:

Methods for site directed mutagenesis for creating *RL*3841 $wreU_{3841}$, $wreQ_{3841}$ mutants, *R. etli* CE3 O-antigen ligase (oal_{PB} *and* oal_{chr}) mutants, complementation of $wreU_{3841}$, $wreQ_{3841}$ mutants with *R. etli wreU* and *wreQ* genes, SDS-PAGE and sequence alignment are described in Chapter Six.

Immunoblot analysis of LPS of RL3841 mutants— After SDS-PAGE, the mini-gel were electrotransferred to nitrocellulose. The dried nitrocellulose blots were immuno-stained using anti-3841 rabbit antisera and anti-rabbit immunoglobulin G (IgG) conjugated with alkaline phosphatase (Sigma). Nitrocellulose was incubated in 25 ml of TSG buffer (50 mM Tris, 0.2 M NaCl, 0.5% gelatin (Sigma), pH 7.4) for 10 min. This was replaced with 15 ml of TSG and 50 µl anti-3841 and allowed to incubate overnight on a rotating shaker. Nitrocellulose was rinsed in five changes of 20 ml of TSG for 10 min each, followed by the addition of a final volume of 15 ml of TSG and 10 µl of secondary antibody and incubation on a rotating shaker for at least 3 hrs. Then, nitrocellulose was rinsed four times with 20 ml of TSG for five min each and developed with 50 mls of AMPD (100 mM 2-amino-2-methyl-1,3-proanediol (AMPD), 1 mM $MgSO_4$, 1 mM levamisole, pH 9.2) containing 0.1 mg/ml of nitrobluetetrazolium (NBT, Fischer Scientific) dissolved in water and 0.125 mg/ml of 5-bromo-4-chloro-3-indolyl phosphate (BCIP, Fischer Scientific) dissolved in N,N-dimethyl-formamide (DMF).