1-1-2017

Synthesis of *N*-Acetyl-□-quinovosamine in *Rhizobium etli* CE3 is Completed After Its 4-keto-precursor is Linked to a Carrier Lipid

Tiezheng Li
Marquette University

K. Dale Noel
Marquette University, dale.noel@marquette.edu

Accepted version. *Microbiology*, Vol. 163, No. 12 (2017). DOI.

Microbiology, Vol. 163, No. 12 (2017): 1890-1901. DOI.

Synthesis of *N*-acetyl-d-quinovosamine in *Rhizobium etli* CE3 is completed after its 4-keto-precursor is linked to a carrier lipid

Tiezheng Li
Department of Biological Sciences, Marquette University, Milwaukee, WI 53233, USA

K. Dale Noel
Department of Biological Sciences, Marquette University, Milwaukee, WI 53233, USA

ABSTRACT

Bacterial O-antigens are synthesized on lipid carriers before being transferred to lipopolysaccharide core structures. *Rhizobium etli* CE3 lipopolysaccharide is a model for understanding O-antigen biological function. CE3 O-antigen structure and genetics are known. However, proposed enzymology for CE3 O-antigen synthesis has been examined very little *in vitro*, and even the sugar added to begin the synthesis is uncertain. A model based on mutagenesis studies predicts that 2-acetamido-2,6-dideoxy-d-glucose (QuiNAc) is the first O-antigen sugar and that genes *wreV, wreQ* and *wreU* direct QuiNAc synthesis and O-antigen initiation. Previously, synthesis of UDP-QuiNAc was shown to occur *in vitro* with a WreV orthologue (4,6-hexose dehydratase) and WreQ (4-reductase), but the WreQ catalysis in this conventional deoxyhexose-synthesis pathway was very slow. This seeming deficiency was explained in the present study after WreU transferase activity was examined *in vitro.* Results fit the prediction that WreU transfers sugar-1-phosphate to bactoprenyl phosphate (BpP) to initiate O-antigen synthesis. Interestingly, WreU demonstrated much higher activity using the product of the WreV

catalysis [UDP-4-keto-6-deoxy-GlcNAc (UDP-KdgNAc)] as the sugar-phosphate donor than using UDP-QuiNAc. Furthermore, the WreQ catalysis with WreU-generated BpPP-KdgNAc as the substrate was orders of magnitude faster than with UDP-KdgNAc. The inferred product BpPP-QuiNAc reacted as an acceptor substrate in an *in vitro* assay for addition of the second O-antigen sugar, mannose. These results imply a novel pathway for 6-deoxyhexose synthesis that may be commonly utilized by bacteria when QuiNAc is the first sugar of a polysaccharide or oligosaccharide repeat unit: UDP-GlcNAc → UDP-KdgNAc → BpPP-KdgNAc → BpPP-QuiNAc.

KEYWORDS

O-antigen, biosynthesis, Rhizobium, bactoprenyl-phosphate, quinovosamine, deoxysugar

ABBREVIATIONS

BpP, bactoprenyl phosphate; BpPP, bactoprenyl pyrophosphate; DGK, kinase encoded by Streptococcus mutans dgk, with preference for polyprenol substrates; GlcNAc, 2-acetamido-2-deoxy-d-glucose, also known as N-acetyl-d-glucosamine; GT, glycosyltransferase; KdgNAc, 2-acetamido-2,6-dideoxy-d-xylo-4-hexulose, also known as 4-keto-6-deoxy-GlcNAc; LPS, lipopolysaccharide; Man, mannose; QuiNAc, 2-acetamido-2,6-dideoxy-d-glucose, also known as N-acetyl-d-quinovosamine; Und-P, undecaprenyl phosphate; Und-PP, undecaprenyl pyrophosphate.

INTRODUCTION

O polysaccharide, or O-antigen, is the outermost component of the lipopolysaccharide (LPS) that is the major constituent of the outer leaflet of the outer membrane in bacteria [1]. Bacterial mutants lacking O-antigen have deficiencies that reveal the profound physiological and ecological importance of this portion of LPS [2–8]. For instance, complete and abundant O-antigen (Fig. 1a) of the model bacterium of this study, *Rhizobium etli* strain CE3, is indispensable for infection and development of nitrogen-fixing root nodules on its legume host, *Phaseolus vulgaris* [9–11].

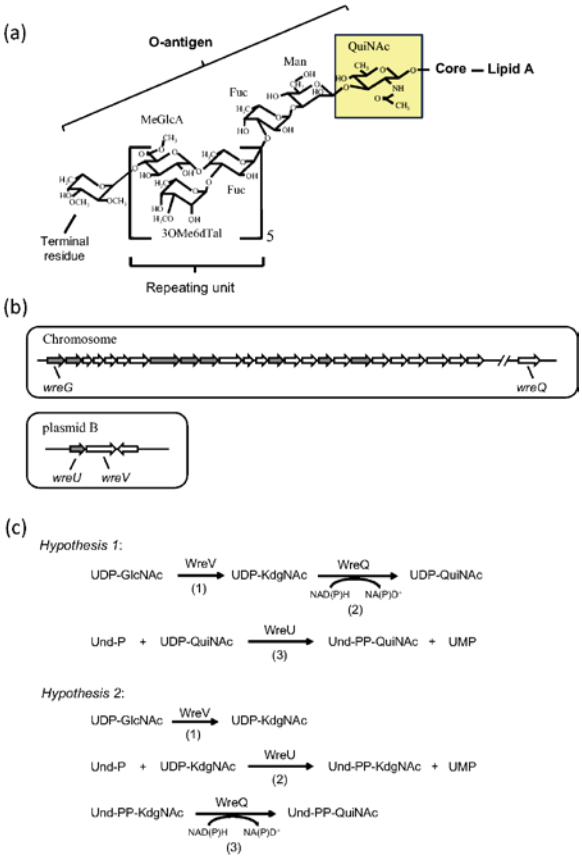


Fig. 1.
(a) Structure of *Rhizobium etli* CE3 O-antigen. The O-antigen structure of *R. etli* LPS is shown linked to the lipid A core. Abbreviations for the sugars: QuiNAc, *N*-acetyl- d-quinovosamine; Man, mannose; Fuc, fucose; MeGlcA, methyl-glucuronate; 3OMe6dTal, 3-O-methyl-6-deoxytalose; terminal residue, TOMFuc, 2,3,4-tri-O-methylfucose or DOMFuc, 2,3-di-O-methylfucose. The proposed first O-antigen sugar, QuiNAc, is highlighted. (b) *R. etli* CE3 O-antigen genetic clusters. Upper panel: the chromosomal *wre* gene cluster (previously called *lps* region α) spanning nucleotides 784 527 to 812 262 of the genome sequence consists of 25 predicted ORFs. Another chromosomal ORF (*wreQ*) spanning nucleotides 2 969 313 to 2 970 242 is required for QuiNAc synthesis [27]. Lower panel: a 4-kilobase cluster on plasmid pCFN42b consists of three predicted ORFs. In each panel the predicted GTase-encoding genes are in grey. Genes encoding enzymes studied in the current work, *wreG*, *wreQ*, *wreU* and *wreV*, are specifically labelled. (c) Two hypotheses of O-antigen initiation in *R. etli* CE3. Reactions in each hypothesis and the enzyme that catalyses each reaction are indicated. In both hypotheses, the first reaction (1) is the same, the conversion of UDP-GlcNAc to UDP-KdgNAc catalysed by the predicted 4,6-dehydratase WreV. The two hypotheses differ in reactions (2) and (3). In hypothesis 1, KdgNAc is reduced to QuiNAc on the UDP linkage by WreQ and then QuiNAc-1-P is transferred by WreU. In hypothesis 2, KdgNAc-1-P is transferred by WreU first and then KdgNAc is reduced to QuiNAc by WreQ on the Und-PP linkage.

The CE3 O-antigen is also an intriguing model for studying polysaccharide biosynthesis. It has features, such as its precisely controlled number of repeat units [12], which are not readily explained by known mechanisms. Making it attractive is also the fact that all 29 genes considered necessary specifically for its synthesis have been mutated. For instance, it has been possible to identify nine genes encoding glycosyltransferases (GTs), and, by biochemical analysis of truncated LPS from each GT mutant, to explain which sugar linkages are catalysed by each of the nine GTs [13] (Fig. 1b). However, these assignments have not been confirmed by investigation *in vitro* with purified enzymes and defined substrates. The initial step of the biosynthesis is a logical first reaction to investigate.

The biosynthesis of all characterized O-antigens is believed to share a conserved initial type of reaction catalysed by a family of GTs that are integral membrane proteins. The initiating GTs catalyse transfer of a sugar-1-phosphate moiety from a nucleotide-sugar donor to the membrane lipid carrier bactoprenyl phosphate (BpP), resulting in bactoprenyl-pyrophosphoryl-sugar (BpPP-sugar) [14, 15]. Due to the difficulty in obtaining purified enzymes and the limited availability of substrates, this initial step in the synthesis of an O-antigen has been demonstrated in only a few cases [16–18]. In *R. etli* CE3 O-antigen (Fig. 1a), the proposed first sugar is 2-acetamido-2,6-dideoxy- d-glucose (d-QuiNAc, hereafter referred to as QuiNAc) [12, 13, 19]. Although QuiNAc is found in a number of bacterial polysaccharides, the mechanism of its incorporation into a polysaccharide, in particular as the initiating sugar, has not been reported. The predicted initiating GT for *R. etli* CE3 O-antigen synthesis is encoded by the *wreU* gene (Fig. 1b) [13]. The LPS of a *wreU* null mutant lacks all O-antigen-specific sugars including QuiNAc [13].

QuiNAc is derived from the central metabolite UDP-GlcNAc (UDP- *N*-acetyl- d-glucosamine [UDP-2-acetamido-2-deoxy- d-glucose]) [19, 20]. Besides WreU, two additional enzyme activities are expected in a pathway from UDP-GlcNAc to BpPP-QuiNAc (Fig. 1c). The first is a 4,6-dehydratase that catalyses conversion of UDP-GlcNAc to UDP-2-acetamido-2,6-dideoxy- d-xylo-4-hexulose (also known as UDP-4-keto-6-deoxyGlcNAc and hereafter referred to as UDP-KdgNAc). *R. etli* gene *wreV* (Fig. 1b) encodes a protein whose predicted sequence aligns with enzymes known to catalyse this reaction *in vitro* [19, 21–25]. The gene for one of these characterized enzymes, *Pseudomonas aeruginosa wbpM*, complements *R. etli wreV* mutants [19].

The other expected enzyme activity is a 4-reductase that catalyses the reduction of the KdgNAc moiety to QuiNAc (Fig. 1c). Analysis of the *R. etli* CFN42 total nucleotide sequence assigns this type of activity to the protein encoded by *wreQ* (Fig. 1b), and mutation of this gene has previously been shown to cause the absence of QuiNAc from the *R. etli* CE3 LPS [26]. Recently, it was shown that WreQ catalyses the conversion of UDP-

KdgNAc to UDP-QuiNAc *in vitro* [19]. However, the catalysis by WreQ was relatively very slow, raising the question of whether UDP-KdgNAc is the natural substrate of WreQ *in vivo*. Also relevant is the fact that in the small amount of LPS O-antigen produced by a *wreQ* null mutant, the QuiNAc residue is replaced by KdgNAc [27]. This result raises the possibility that WreU acts on either the QuiNAc or KdgNAc moiety, or, considering the observed slow WreQ catalysis with UDP-KdgNAc as the substrate, the normal route to QuiNAc *in vivo* may be the one shown in Fig. 1(c) as hypothesis 2.

In the present study, the alternative hypothetical pathways of Fig. 1(c) were tested by an *in vitro* biochemical approach using enzymes expressed from hybrid cloned genes in *Escherichia coli*. In addition, BpPP-QuiNAc was shown to be a functional acceptor substrate for the next step in CE3 O-antigen synthesis in an assay *in vitro* using the predicted sugar donor and the predicted transferase encoded by gene *wreG*.

RESULTS

Recombinant *R. etli* WreU was expressed in *E. coli*

The *wreU* gene of *R. etli* CE3 was cloned into a pET15b vector, yielding a genetic construct from which the expressed WreU protein included an amino-terminal six-histidine (His_6) tag. When this ORF was subcloned into a vector that replicates in *R. etli,* its expression complemented the LPS-deficient phenotype of *R. etli wreU*-null mutant strain CE566 (Fig. S1, available with the online version of this article). After overexpression in *E. coli*, His_6-WreU was found exclusively in the cell membrane fraction. Attempts to purify His_6-WreU free of membrane were not successful despite trying various detergents, various expression conditions, and making other types of WreU constructs. Thus, the *E. coli* membrane fraction containing His_6-WreU was used in the *in vitro* studies of WreU.

WreU possessed GT activity with preference for UDP-KdgNAc as the nucleotide-sugar substrate

For testing WreU enzymatic activity, the lipid carrier substrate, undecaprenyl phosphate (Und-P), was synthesized *in situ* from undecaprenol and ATP with an enzyme having polyprenol kinase activity as described by [28]. UDP-KdgNAc or UDP-QuiNAc, each enzymatically synthesized as described previously [19], or UDP-GlcNAc was added as a possible nucleotide-sugar substrate. The reactions were started by adding WreU-containing membranes, or control membranes lacking WreU, and terminated by chloroform-methanol extraction. Und-PP-sugars, such as the predicted products of WreU catalysis, partition into the organic phase in this type of extraction thereby separating them from the nucleotide-sugar substrates [29, 30].

To facilitate visualization of the product after TLC separation, the lipid substrate Und-P was labelled with ^{32}P by using ATP (γ-^{32}P) in its synthesis. The result of the WreU assay is shown in Fig. 2 . An abundant product corresponding to an undecaprenyl pyrophosphate-linked sugar (Und-PP-sugar) candidate (compound I) was detected in the reaction only with UDP-KdgNAc as the nucleotide-sugar substrate (Fig. 2 , lane 4). In the reaction with an equal concentration of UDP-QuiNAc, a faint spot representing a different Und-PP-sugar candidate (compound II) was observed (Fig. 2 , lane 5), and no product was detected in the reaction with UDP-GlcNAc (Fig. 2 , lane 3). When quantified with a phosphorimager, compound I in lane 4 had a 30-fold higher intensity than compound II in lane 5. Importantly, the production of compounds I and II required both the lipid substrate Und-P and the enzyme WreU (Fig. 2 , lane 1, 2). These results provided evidence that WreU is an initiating GT and UDP-KdgNAc is the preferred sugar-P-donor substrate. UDP-QuiNAc was much less favoured, and UDP-GlcNAc led to no detectable reaction.

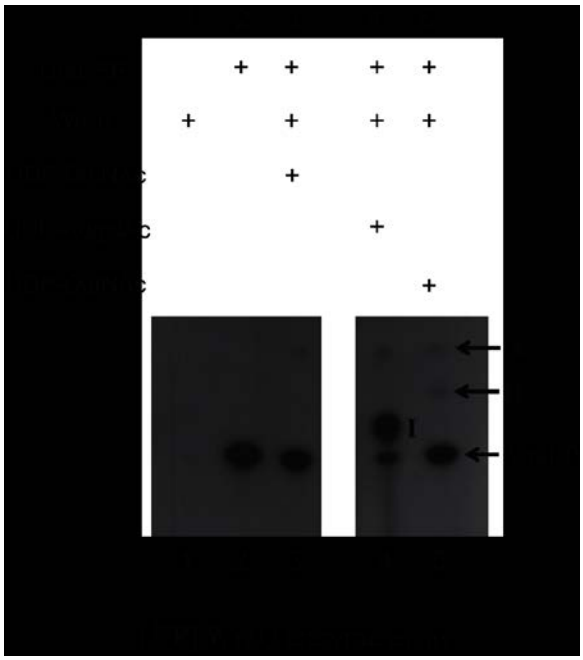


Fig. 2.
Nucleotide sugar substrate specificity of *R. etli* WreU and TLC analysis of the products of its activity. The images show TLC separation of products extracted into the organic phase after generation in WreU reactions that included ^{32}P-labelled lipid substrate Und-P. Two portions of one autoradiogram from one experiment are shown. Lanes: lane 1, by omitting the polyprenyl kinase for its synthesis, lipid substrate Und-P was absent (negative control 1); lane 2, no WreU crude enzyme was added to the reaction (negative control 2); lanes 3–5, different nucleotide sugar substrates were added, as indicated in the table above the TLC images. Inferred compounds: I, Und-PP-KdgNAc; II, Und-PP-QuiNAc; *, a compound derived from Und-P in the presence of *E. coli* membrane (its Rf value with a different TLC solvent (not shown) matches that of Und-PP [44]).

WreQ catalysed the reduction of KdgNAc to QuiNAc on Und-PP linkage

The foregoing results with WreU were consistent with the second step of hypothesis 2 (Fig. 1c). Hence, the next step of this hypothesis was tested: does WreQ catalyse the reduction of Und-PP-KdgNAc to Und-PP-QuiNAc? His$_6$-WreQ had been produced and purified in a previous study, in which it had been shown to catalyse the reduction of d-KdgNAc stereospecifically to d-QuiNAc [19].

A WreU-WreQ-coupled assay was carried out with ^{32}P-radiolabelling (Fig. 3). The Und-^{32}PP-KdgNAc (compound I) produced in the WreU reaction (Fig. 3 , lane 1) served as a substrate for WreQ. NADH was chosen as the reducing substrate. 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. 3 , lane 4). When NADH alone was added without WreQ, no conversion occurred (Fig. 3 , lane 3). When WreQ was added but NADH was omitted, a very small amount of compound II was produced (Fig. 3 , lane 2), possibly due to contaminating NADH from the crude WreU enzyme (membrane).

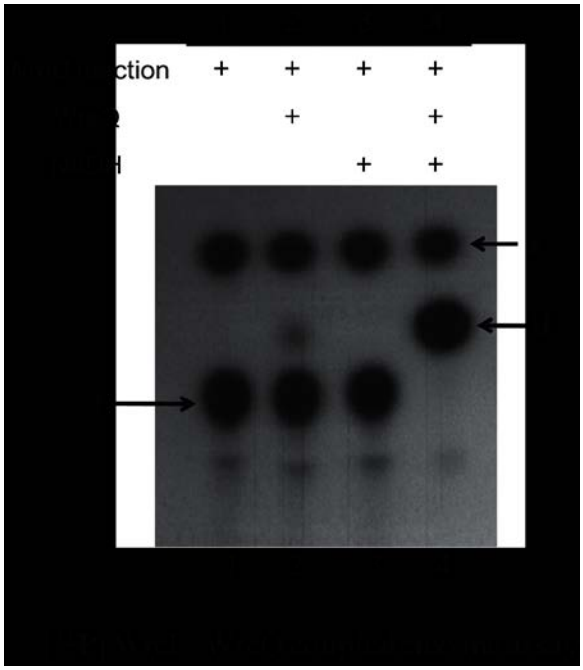


Fig. 3.
Alteration of WreU reaction product by addition of WreQ and NADH. The autoradiogram shows TLC separation of products extracted into the organic phase after generation in WreU-WreQ-coupled enzyme assays. All reactions were set up as noted for lane 4 of Fig. 2 , with UDP-KdgNAc as the nucleotide substrate for WreU, and the reactions were allowed to proceed for 1 h. The reactions then varied by whether WreQ and NADH were added at this point, and incubation continued for another hour. 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. Inferred compounds: I, Und-PP-KdgNAc; II, Und-PP-QuiNAc; *, Und-PP (see Fig. 2 legend).

WreQ catalysis was much faster when KdgNAc was linked to Und-PP rather than UDP

The result of the WreU-WreQ-coupled assay suggested that WreQ has Und-PP-KdgNAc (compound I) reductase activity, leading to Und-PP-QuiNAc (compound II) as the product. WreQ can also catalyse UDP-KdgNAc reduction to UDP-QuiNAc *in vitro*, but that reaction is relatively slow [19]. The rates of catalysis with Und-PP-KdgNAc as the substrate versus UDP-KdgNAc as the substrate were compared by TLC and autoradiography in the following experiments.

To provide conditions for estimating the rate of a WreQ-catalysed reduction of the lipidated substrate Und-PP-KdgNAc, the WreU-WreQ-coupled reactions were performed with serially diluted WreQ concentrations, and the WreQ reaction was allowed to proceed for only 1 min (instead of 1 h in the experiment shown in Fig. 3). The conversion of Und-PP-KdgNAc (compound I) to Und-PP-QuiNAc (compound II) gradually increased with decreasing dilution of WreQ (from 10^{-6} to 10^{-2}) (Fig. 4a , lanes 2–6). The negative control (0 min) indicated that the method to stop the reaction was effective (Fig. 4a , lane 1). At 10^{-2} dilution, the conversion was almost complete in 1 min (Fig. 4a , lane 6), whereas the reactions with 10^{-3} and 10^{-4} diluted WreQ enzyme were slow enough that reaction rates/enzyme concentration (V/[E]) could be estimated (Table 1).

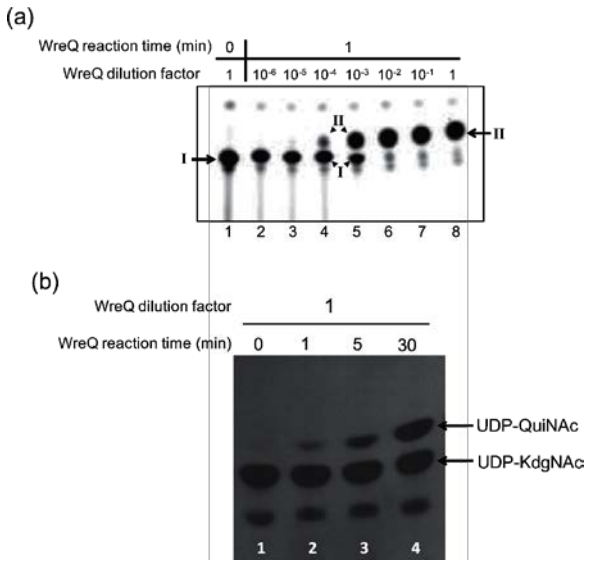


Fig. 4.
Comparison of the WreQ kinetic activity with lipid-linked versus nucleotide-linked substrates. Autoradiograms show products separated on TLCs after controlled times of incubation and concentrations of WreQ. (a) WreQ catalysis with the lipidated substrate. First, [^{32}P] WreU transferase reactions with the UDP-KdgNAc substrate were allowed to proceed for 1 h (i.e. under the same conditions as for Fig. 2 , lane 4). Then, WreQ and NADH were added. After 1 min, the reactions were terminated by adding and rapidly mixing with chloroform-methnol/3 : 2 (solvent I). Lanes: lane 1 is a 0 min control in which solvent I was added before WreQ; lanes 2–8 are groups that contain serially diluted WreQ, from 10 ^{6} to 1 (undiluted). Inferred compounds: I, Und-PP-KdgNAc; II, Und-PP-QuiNAc. (b) WreQ catalysis with UDP-KdgNAc. WreQ and NADH were added to reactions in which UDP-[^{3}H]KdgNAc was produced by complete conversion from UDP-[^{3}H]GlcNAc catalysed by WbpM in a 30 min reaction [19]. WreQ concentration was the same as the undiluted concentration used in panel (a) (10 µg His $_{6}$-WreQ per 100 µl reaction). Catalysis was allowed to proceed for different times before being terminated by boiling. Lanes: lane 1, 0 min; lane 2, 1 min; lane 3, 5 min; lane 4, 30 min.

Table 1. Comparing the estimated enzymatic activities of WreQ with the two substrates, UDP-KdgNAc and Und-PP-KdgNAc

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−1)	Substrate conversion per active site per min *V*/[*E*] (min−1)
UDP-KdgNAc	500	28	45	225	30	7.5	0.27
Und-PP-KdgNAc	0.33	0.0028	62.5	0.206	1	0.206	74
	0.3	0.00028	14	0.042	1	0.042	150

Substrate conversion was calculated from the TLC result shown in Fig. 4 . The intensities of spots were measured by ImageQuantTL software for ^{32}P spots in Fig. 4(a) and by ImageJ software for ^{3}H spots in Fig. 4(b) .

For visual comparison, the WreQ catalysis with the nucleotide substrate UDP-KdgNAc was carried out with a single (much higher) WreQ concentration and terminated at different time points (Fig. 4b). The reaction showed near linear progression and the data obtained at 30 min reaction time (Fig. 4b , lane 2) was used for calculation of the reaction rate (Table 1). The *V*/[*E*] calculated from the results in Fig. 4 indicated that the WreQ catalysis was at least two orders of magnitude faster when KdgNAc was linked to Und-PP rather than UDP (Table 1), even though the lipidated substrate was presented at 1000-fold lower concentration.

WreG catalysed the addition of the second O-antigen sugar (mannose) to Und-PP-QuiNAc

A previous study based on mutant phenotypes proposed that WreG is the GT that transfers the second O-antigen sugar mannose (Man) to QuiNAc [13]. The *in vitro* assay system developed in the present study provided a means to obtain biochemical evidence for the role of WreG and to confirm that Und-PP-QuiNAc is the precursor for further O-antigen synthesis. As a first step, the *wreG* gene of *R. etli* CE3 was cloned into vector pET21b such that the expressed WreG protein in *E. coli* included a carboxy-terminal His $_6$ tag. When this ORF was subcloned into a vector that replicates in *R. etli,* its expression complemented the LPS-deficient phenotype of *R. etli wreG*-null mutant strain CE358 (Fig. S2). To test the enzymatic activity of WreG-His $_6$, two potential acceptor substrates, Und-PP-KdgNAc (compound I) and Und-PP-QuiNAc (compound II), were produced by the WreU reaction and the WreU-WreQ-coupled reaction, respectively (Fig. 5 , lanes 2 and 3). GDP-Man was added as the donor of Man, and WreG was added as a crude membrane fraction (Fig. 5 , lanes 4 and 5) or as purified enzyme (Fig. 5 , lane 6).

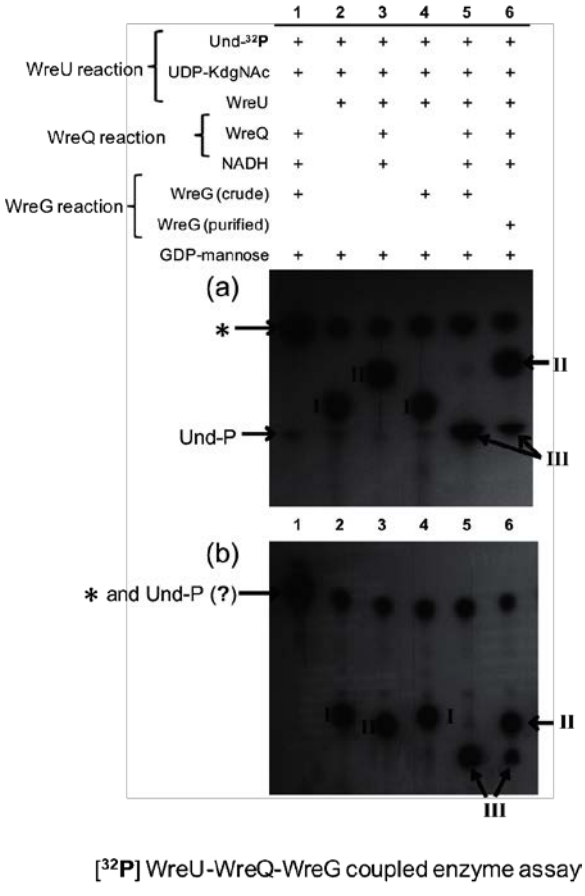


Fig. 5.

The product generated by the combined action of WreU, WreQ and predicted Bp-PP-QuiNAc mannosyl transferase WreG. ^{32}P-labelled compounds were extracted into the organic phase after reactions with six varied

combinations of WreU, WreQ, WreG and substrates. The two panels show autoradiograms of TLCs carried out with two different solvents, 2-propanol/ammonium hydroxide/water, 6 : 3 : 1 (a) and chloroform/methanol/water, 65 : 25 : 4 (b). All reaction mixtures contained substrates UDP-KdgNAc, Und-^{32}P and GDP-Man. They varied by lacking added NADH or enzyme as follows: lane 1, no WreU; lane 2, no WreQ, WreG or NADH; lane 3, no WreG; lane 4, no WreQ or NADH; lane 5, WreG activity provided by crude membrane fraction and all other reagents and enzymes added; lane 6, all reagents and enzymes added, including purified WreG enzyme. Inferred compounds: I, Und-PP-KdgNAc; compound II, Und-PP-QuiNAc; compound III, Und-PP-QuiNAc-Man; *, Und-PP (see Fig. 2 legend).

Und-PP-KdgNAc (compound I) remained almost unchanged when crude WreG and GDP-Man were included in the WreU reaction (compare lane 4 to lane 2 in Fig. 5). In contrast, the Und-PP-QuiNAc (compound II) produced when WreQ was added to the WreU reaction mixture was converted to a slower-moving compound (compound III) when WreG and GDP-Man were also added (as shown by lanes 5 and 6 compared to lane 3 in Fig. 5). When WreG was added as a crude membrane preparation, compound II was almost completely replaced by compound III (Fig. 5 , lane 5), whereas the purified WreG was less active, as revealed by partial conversion of compound II to compound III in lane 6 of Fig. 5 .

Based on the substrate requirements for its formation, compound III was inferred to be Und-PP-QuiNAc-Man, the predicted lipid-linked disaccharide resulting from the GT activity of WreG. Because compound III migration was nearly identical to Und-P in TLC solvent A (Fig. 5 a), a different solvent (solvent B) was also used. This second solvent system separated compound III from Und-P and other radiolabelled compounds (Fig. 5 b).

The result of this assay provided *in vitro* evidence that WreG is the mannosyltransferase for adding the second O-antigen sugar (Man) in *R. etli* CE3. Furthermore, Und-PP-QuiNAc was utilized much more readily than Und-PP-KdgNAc as the acceptor of Man *in vitro*, suggesting that normally the reduction of KdgNAc to QuiNAc by WreQ would occur before Man addition.

DISCUSSION

Results in this study lead to the following inferences regarding QuiNAc and O-antigen synthesis in *R. etli*, as discussed further in succeeding paragraphs: (1) the conversion of d-GlcNAc to d-QuiNAc in *R. etli* occurs in an unconventional manner compared with other characterized deoxysugar syntheses (i.e. it follows hypothesis 2 of Fig. 1c). As with biosynthesis of other 6-deoxysugars, it proceeds with formation of a 4-keto-6-deoxyhexose intermediate (KdgNAc). However, the KdgNAc-P intermediate is first transferred from nucleotide linkage to a bactoprenyl phosphate carrier before undergoing 4-reduction to the 6-deoxy product, QuiNAc. (2) The switch from nucleotide linkage to lipid carrier is directed by WreU, whose reaction requirements confirm the prediction that it is the initiating GT for *R. etli* O-antigen biosynthesis. The specificity of WreU for KdgNAc dictates that bactoprenyl-PP-KdgNAc is the first lipid-linked O-antigen intermediate. (3) The unconventional lipidated-substrate specificity of 4-reductase WreQ is responsible for delaying conversion of KdgNAc to QuiNAc until KdgNAc is attached to bactoprenol-PP. Without WreQ, KdgNAc would be the predicted proximal sugar of the final O-antigen, and, in fact such is the case in the small amount of O-antigen produced in *wreQ*-null mutant strain CE166 [27]. (4) The acceptor-substrate specificity of mannosyltransferase WreG ensures that QuiNAc predominately replaces KdgNAc before O-antigen synthesis can proceed. WreG operates fastest with bactoprenyl-PP-QuiNAc and thereby dictates that QuiNAc and WreQ are needed for efficient synthesis of the *R. etli* CE3 O-antigen.

WreU is a UDP-KdgNAc:bactoprenyl-P KdgNAc-1-P transferase (reaction 2 of hypothesis 2 in Fig. 1c). This conclusion is based on (1) the substrate requirements of the reaction it catalyses, (2) physical and chemical properties of the inferred product, and (3) the degree of sequence alignment with other characterized initiating GTs. The lipid substrate of WreU in the *in vitro* reactions was undecaprenyl phosphate (Und-P or C_{55}-P).

However, the exact form of bactoprenol lipid carrier in *R. etli* CE3 is not known. It is likely dodecaprenol-P (C_{60}-P) as reported for *Rhizobium leguminosarum* 3841 and *Sinorhizhobium meliloti* 1021 [31]. Of the two hypothetical sugar-donor substrates (Fig. 1c , hypothesis 1 vs hypothesis 2), UDP-KdgNAc yielded 30-fold greater activity than UDP-QuiNAc *in vitro*. The lack of activity with UDP-GlcNAc as the donor substrate indicates that WreU requires the 6-deoxy moiety for activity. The products, compounds I (from UDP-KdgNAc) and II (from UDP-QuiNAc), behaved in solvent extraction and relative TLC migration as would be predicted. They also carried the input ^{32}P of the lipid substrate. The WreU enzyme assay was also carried out with different radioisotope labelling in which the UDP-KdgNAc carried tritium [^{3}H] in the sugar moiety. After extraction into the organic solvent, the ^{3}H-labelled product showed the same relative migration on TLC as the ^{32}P-compound I (data not shown). Hence, radiolabelling provided additional evidence of both the input lipid and the sugar being present in the product.

A final argument for the enzymatic identity of WreU is based on the predicted amino-acid sequence of translated *wreU*. The predicted topology and sequence alignment of WreU (Fig. S3) places it within a large subgroup of the superfamily of initiating GTs represented by the active carboxy-terminal portion of *Salmonella* WbaP [18], WecP from *Aeromonas hydrophila* AH-3 [17], and PglC from *Campylobacter jejuni* NCTC 11168 [32]. Like other members of this subgroup, WreU is predicted to have a single transmembrane segment near the amino-terminus followed by a cytoplasmic catalytic domain that constitutes the rest of the polypeptide of these 'small' phospho-GTs [33].

WreQ catalysed the inferred 4-reduction of KdgNAc to QuiNAc orders of magnitude faster when KdgNAc was attached to Und-PP than when it was attached to UDP. The very slow reduction of UDP-KdgNAc by the WreQ catalysis reported in a previous study [19] is thereby explained. The study of that slower reaction, however, had the advantage that it was chemically very clean and allowed definitive demonstration that the QuiNAc produced by the WreQ catalysis has the d-stereo configuration [19].

Although this may be the first report of 6-deoxyhexose biosynthesized in this way, a conceptually analogous precedent is *N*-acetylgalactosamine (GalNAc) synthesis in *E. coli* by the Gnu pathway, in which the WecA-initiating GT first transfers GlcNAc-1-P to Und-P and the Und-PP-GlcNAc product is converted to Und-PP-GalNAc via a Gnu epimerase [34, 35]. The overall outcome is to generate a primer for synthesis of a polysaccharide or oligosaccharide repeat that will have GalNAc at its reducing terminus. This is exactly analogous to the apparent metabolic role of the WreV-WreU-WreQ pathway and its product bactoprenyl-PP-QuiNAc in *R. etli*.

Another pathway for d-QuiNAc synthesis was reported recently [20]. It proceeds by the first two steps as outlined in hypothesis 1 of Fig. 1c , i.e. the path *not* followed by WreV-WreU-WreQ in *R. etli*. Discovered in *Bacillus cereus* strain ATCC 14579, it is catalysed by a 4,6-dehydratase and a 4-reductase that are not homologous with WreV and WreQ [20]. Whereas the bactoprenyl-P-coupled pathway of the Proteobacteria seems suited to provide QuiNAc only to begin polysaccharides and oligosaccharide repeat units, this pathway in the bacilli conceivably could be used to provide QuiNAc either for interior glycosyl positions or the initial position in a growing chain. Surprisingly, though, this more conventional pathway may be very limited phylogenetically. The 4-reductase, Preq, shows high sequence similarity only with proteins in other bacilli and perhaps certain closely related firmicutes. An extensive database search did not find it in Proteobacteria, where the bactoprenyl-P-coupled pathway catalysed by WreV-WreU-WreQ orthologues is widely distributed (Table S1).

Results obtained with WreG validated the functionality of the *in vitro* product of WreU and WreQ activity for CE3 O-antigen synthesis. Based on its structure and the responsible *wre* genes, the CE3 O-antigen synthesis has been deduced [13] to follow the lesser known of the two common overall mechanisms of O-antigen synthesis [1], in which the complete O-antigen with all repeat units is made on the cytoplasmic face of the inner membrane and then transported across the membrane. The model for CE3 O-antigen [13] proposes that QuiNAc is the primer residue [1, 36] for the remainder of O-antigen synthesis, with Man being the next

'adaptor' sugar added (Fig. 1a). WreG is the predicted transferase that catalyses this addition and GDP-Man is the donor substrate for Man addition. The results of the *in vitro* assay of WreG activity (Fig. 5) supported both predictions of the model and Und-PP-QuiNAc as the product from the WreU and WreQ catalysis. The existence of WreQ and the specificity of WreG are coupled. Selectivity for the 4-OH of QuiNAc by WreG [i.e. much slower catalysis with bactoprenyl-PP-KdgNAc (Fig. 5)] is the key reason that WreQ-null mutants have low abundance of O-antigen [26]. However, WreG *in vivo* apparently has enough activity with KdgNAc as the Man acceptor that such mutants have a low amount of O-antigen that is identical to the normal O-antigen except for substitution of KdgNAc for QuiNAc [27]. A faint spot visible in lane 4 of Fig. 5 , (circled in Fig. S4a) may be due to this lower activity of WreG with Und-PP-KdgNAc as the acceptor substrate *in vitro*. This logic leads to the prediction that greatly increasing the specific concentration of just the WreG enzyme will lead to higher O-antigen production in a *wreQ*-minus genetic background. Fig. S4(b) shows results that confirm this prediction. This result explains the basis of the genetic suppression of the WreQ-minus phenotype by multiple copies of the main *wre* cluster [26]. Importantly, it also supports hypothesis 2 over hypothesis 1 of Fig. 1(c) by means of *in vivo* results that are independent of the *in vitro* assays.

The phylogenetic distribution of this pathway of QuiNAc synthesis was investigated by blast searches of the sequenced protein database (Table S1). At least 40 genera had at least one strain that carried orthologues of all three genes– *wreV, wreU* and *wreQ*. Two genera of green-sulfur bacteria had strong matches, but almost all of the rest were in the Proteobacteriaceae, with all its subphyla being represented (Table S1). blast e-values were less than e^{-30} for all three homologues in all strains.

It should be noted that WreQ orthologues are the genes needed specifically for QuiNAc synthesis by the bactoprenyl pathway. As stated above, WreV-WreQ are often linked with WreU orthologues, but, at lower frequency, they are found with orthologues of WbpL, another initiating GT. In a limited search of WreQ hits with e-values below e^{-89}, 211 were linked to a WreU homologue (Table S1) and 63 to a WbpL homologue (Table S2). WreU and WbpL are not homologous; they represent the two very different types of initiating GT structures. It is reasonable to suppose that other initiating GT subtypes are coupled with WreVQ in a strain, depending on how the genetic cluster has evolved.

Recently, *Colwellia psychrerythraea* 34 h, was found to make an 'antifreeze' polysaccharide that has a repeat unit containing QuiNAc [37]. Its ability to synthesize QuiNAc had been predicted the previous year [19] because it had WreV and WreQ sequence matches with very low e-values. A gene whose encoded protein has a significant match with WreU is adjacent to the *wreQ* orthologue on the genome, and the *wreV* orthologue is separated by three genes (Fig. S5).

In summary, results in this study strongly support the second of the two alternative hypotheses of Fig. 1(c). A main conclusion is that biosynthesis of QuiNAc in *R. etli* CE3 (and probably in many other bacteria) is tightly coupled to initiation of the synthesis of a polysaccharide on which it is ultimately the first (reducing-end) sugar. Fig. 6 depicts this coupling, the steps in the pathway, and its association with the membrane. Phylogenomic searches suggest that this pathway is distributed widely among the Proteobacteria. The outcome is Bp-PP-QuiNAc, which in *R. etli* CE3 becomes the platform for the rest of O-antigen synthesis, the next step of which was also demonstrated in this study and is depicted in Fig. 6 as well.

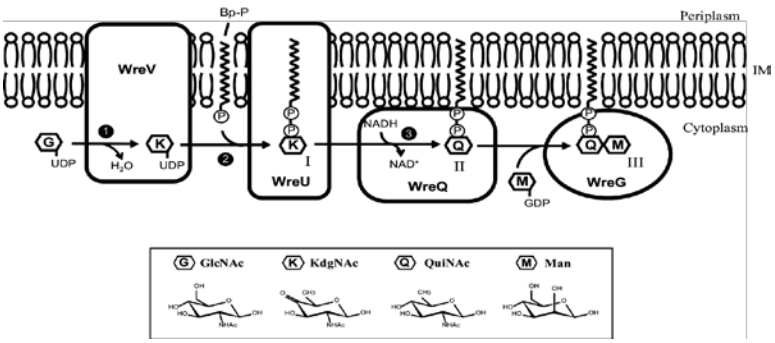


Fig. 6.
Model of QuiNAc synthesis coordinated with O-antigen initiation. The three phases of QuiNAc synthesis in *R. etli* CE3 are indicated with numbers: first, UDP-GlcNAc is converted to intermediate UDP-KdgNAc by the dehydratase WreV (phase 1); second, KdgNAc-1-P is transferred by WreU to the bactoprenyl phosphate (BpP) lipid carrier (phase 2); third and last, the KdgNAc moiety is reduced to QuiNAc by WreQ (phase 3). The final product of QuiNAc synthesis serves as the platform for further O-antigen synthesis to which a Man was transferred by WreG. Compounds: I, Bp-PP-KdgNAc; II, Bp-PP-QuiNAc; III, Bp-PP-QuiNAc-Man.

METHODS

Bacterial strains and growth conditions

Rhizobium etli CE3 was derived from *R. etli* wild-type strain CFN42 by a spontaneous mutation conferring resistance to streptomycin [38]. As in almost all past 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 $CaCl_2$]. 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).

DNA techniques

Genomic DNA was isolated using GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich) and plasmid DNA was isolated using QIAprep Spin Miniprep Kit (Qiagen). DNA extraction from agarose gels was performed using Gel/PCR DNA Fragments Extraction Kit (IBI Scientific). DNA amplification by PCR was performed using Expand High Fidelity PCR System (Roche Applied Science). Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (NEB).

Cloning of *R. etli* wreU and wreG for overexpression

The *R. etli* CE3 *wreU* gene was amplified from *R. etli* CE3 genomic DNA using primers 5′-CCGG CATATGGGCTTGAAACGGGCG-3′ (forward) and 5′-GGCC GGATCCCTAGTGCTTTATTCC-3′ (reverse). The PCR product was cloned into the pET15b vector (Novagen) using NdeI and BamHI sites, generating plasmid pLS5. It encodes the WreU protein with additional amino acids at the amino-terminus, MGSS HHHHHHSSGLVPRGSH (the 6xHis tag is underlined). This WreU construct is referred to as His_6-WreU in this work.

The *R. etli* CE3 *wreG* gene was amplified from *R. etli* CE3 genomic DNA using primers 5′- GC GCTAGCATGAGAGTCCTTCATTT-3′ (forward) and 5′-TT CTCGAGGCGGGAACCGGCCACGT-3′ (reverse). The PCR product was cloned into the pET21b vector (Novagen) using NheI and BamHI sites, generating plasmid pTL59. It encodes the WreG protein with amino-terminal additional amino acids MAS, and carboxy-terminal additional amino acids LE HHHHHH (the 6xHis tag is underlined). This WreG construct is referred to as WreG-His_6 in this work.

Overexpression of His $_6$-WreU, WreG-His $_6$ and the polyprenyl kinase (DGK)

Plasmid pLS5 (His $_6$-WreU) and pTL59 (WreG-His $_6$) were separately transformed into *E. coli* BL21(DE3) cells by electroporation. The polyprenol kinase used in these studies is expressed from cloned *dgk* DNA from *Streptococcus mutans*. Although homologous to *E. coli dgk*, the protein encoded by *Streptococcus mutans dgk* has higher [28], or much higher [39], kinase activity with undecaprenol as the substrate than with diacylglycerols. BL21 cells carrying a pET vector construct encoding this protein with a carboxy-terminal His $_6$-tag [28] was provided by Dr Barbara Imperiali, Massachusetts Institute of Technology, Cambridge, MA. Hereafter in this section, it will be referred to as DGK, to conform with the extant abbreviation in the literature.

Expression of His $_6$-WreU, WreG-His $_6$ and DGK followed the same procedure: a flask of 1 l LB medium containing appropriate antibiotics (ampicillin 100 µg ml $^{-1}$ for His $_6$-WreU and WreG-His $_6$, kanamycin 50 µg ml $^{-1}$ for DGK) was inoculated with a 5 ml overnight start culture and shaken at 37 °C until an optical density between 0.6 and 0.8 was reached. Then the flask was chilled for 1 h. Protein expression was induced by adding IPTG to the culture (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 h 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.

Complementation of *R. etli* CE3 mutants with the respective His-tagged constructs

The DNA sequence-encoding His $_6$-WreU together with the RBS sequence was amplified from plasmid pLS5 with primers 5′-GCC GAATTCATACCCACGCCGAAACAAG-3′ (forward) and 5′-GCC GGTACCAGTTCCTCCTTTCAGCAAA-3′ (reverse). The PCR product was cloned into plasmid pFAJ1708 [40] with EcoRI and KpnI sites, generating plasmid pLS22.

The DNA sequence-encoding WreG-His $_6$ together with the RBS sequence was amplified with primers 5′-GCC GAA TTCATACCCACGCCGAAACAAG-3′ (forward) and 5′-GCC GGTACCAGTTCCTCCTTTCAGCAAA-3′ (reverse). The PCR product was cloned into plasmid pFAJ1708 [40] with XbaI and KpnI sites, generating plasmid pTL63.

Separately, pLS22 (*His $_6$-wreU*) was transferred into CE566 (*wreU*::Km) and pTL63 (*wreG-His $_6$*) was transferred into CE358 (*wreG*::Tn *5*) by triparental mating [41] with plasmid-mobilizer strain MT616 [42], as described previously [13]. Strains containing these constructs were selected on TY agar plates supplemented with 200 µg of streptomycin ml $^{-1}$, 30 µg of nalidixic acid ml $^{-1}$, 5 µg tetracycline ml $^{-1}$, 30 µg of kanamycin ml $^{-1}$. Single colonies were purified and analysed by SDS-PAGE.

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. Frozen cell pellets from 500 ml culture were thawed with lysis buffer (buffer A for DGK, 50 mM Tris, 1 mM ethylenediaminetetraacetic acid; buffer B for His $_6$-WreU, 20 mM Tris, 300 mM NaCl, pH 8.5; buffer C for WreG-His $_6$, 20 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, pH=7.0, with 14.3 mM 2-mercaptoethanol), and lysed by sonication. The lysate was centrifuged first at a low speed (6000 ***g*** , 20 min at 4 °C) to remove most of the cellular debris and then followed by a high speed spin (65 000 ***g*** , 120 min at 4 °C) to pellet the cell membrane fraction (stored at −80 °C if not used). The pellet of His $_6$-WreU and WreG-His $_6$ was homogenized in the respective lysis buffer and aliquoted into 100 µl fractions for storage at −80 °C.

Purification of DGK from membrane fractions

Frozen cell membrane was thawed and resuspended in 0.5 ml buffer D (20 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, pH 8.0) and incubated with 1 % CHAPS for 1 h 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 D 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 dialysed and concentrated with an 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

One tube of 0.5 ml frozen cell membrane fraction was thawed and incubated with 1 % Triton X-100 for 2.5 h 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 C 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.

In vitro enzyme assays

WreU GT assay – the lipid substrate Und-P was prepared according to [28] with modification. Briefly, 3 µl DMSO and 10 µl 10 % Triton X-100 were mixed with 13 nmol of dried undecaprenol (American Radiolabeled Chemicals). The tube was vortexed to ensure solubilization of the lipid. To the same tube, 5 µM [γ- ^{32}P]-ATP (2000 mCi mmol $^{-1}$) (PerkinElmer), 1 µl of purified DGK (~50 ng), 50 mM Tris buffer, pH 8.0, 40 mM $MgCl_2$ were added to a total volume of 100 µl. The DGK reaction was incubated at 30 °C for 1 h. To start the WreU enzyme assay, 1 µl (~2 µg) His $_6$-WreU membrane fraction was added to the DGK reaction. In the [^{32}P]-WreU assay, nucleotide sugar substrates tested were: UDP-GlcNAc (Sigma), purified UDP-KdgNAc and UDP-QuiNAc. The concentration of each nucleotide sugar substrate was 0.05 mM.

The WreU reactions were incubated at 30 °C for 1 h, 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) [43]. 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 (2-propanol/ammonium hydroxide/water, 6 : 3 : 1). Dried TLC plates were exposed to films or photostimulable phosphor (PSP) plates and viewed by autoradiogram.

WreU-WreQ-coupled assay – the WreU reactions with UDP-KdgNAc as the substrate were incubated for 1 h at 30 °C. To the WreU reactions, 0.1 mM NADH and 1 µl (~10 µg) WreQ enzyme were added. The reactions were allowed to proceed for 1 h at 30 °C after WreQ addition. Then they were quenched and prepared for analysis as described above for WreU enzyme assay.

WreU-WreQ-WreG-coupled reaction – in reactions that aimed to test the GT activity of WreG, 10 µl crude (~10 µg) or purified WreG enzyme (~40 µg) and 0.1 mM GDP-Man were added to [^{32}P]-WreU reactions together with (or without) 10 µg WreQ and 0.1 mM NADH. Reactions were allowed for 1 h after adding WreG and then quenched and prepared for TLC analysis. For analysis of the reaction products, two TLC solvents were used: solvent A and solvent B (choloform/methanol/water, 65 : 25 : 4).

Rate comparison of WreQ-catalysed reaction with different substrates

To estimate the rate of the WreQ reaction with nucleotide sugar substrate, 10 µg His $_6$-WreQ protein and 1 mM NADH were added to the WbpM reaction which was incubated for 30 min to generate product UDP-KdgNAc [19] and the WreQ reactions were quenched at 1, 5 and 30 min. Reducing the WreQ enzyme concentration by a

factor of 10 was attempted which led to a very slow reaction, thus only one concentration of WreQ was used for this reaction.

To estimate the rate of the WreQ reaction with the lipidated substrate, firstly [^{32}P] WreU transferase reactions with the UDP-KdgNAc substrate for 1 h. Then 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 and the WreQ reactions were allowed for only 1 min. In one reaction, solvent I was added before the addition of the WreQ enzyme as a 0 min control, to show that the method of quenching the reactions was effective. The organic phases of these reactions were analysed by TLC. Radioactive (^{32}P) spots were quantified by phosphorimager and used for estimating reaction kinetics.

FUNDING INFORMATION

This work was supported by National Institutes of Health Grant 1 R15 GM087699-01A1.

ACKNOWLEDGEMENTS

We thank Dr Barbara Imperiali for the gift of an expression vector construct encoding the *Streptococcus mutans* undecaprenol kinase with a carboxy-terminal His $_6$-tag and Dr J. S. Lam for providing the WbpM-His-S262 expression vector.

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest.

REFERENCES

1. Raetz CR, Whitfield C. Lipopolysaccharide endotoxins. Annu Rev Biochem 2002; 71: 635– 700
2. Frank MM, Joiner K, Hammer C. The function of antibody and complement in the lysis of bacteria. Rev Infect Dis 1987; 9: S537– S545
3. Bowden MG, Kaplan HB. The Myxococcus xanthus lipopolysaccharide O-antigen is required for social motility and multicellular development. Mol Microbiol 1998; 30: 275– 284
4. Toguchi A, Siano M, Burkart M, Harshey RM. Genetics of swarming motility in Salmonella enterica serovar typhimurium: critical role for lipopolysaccharide. J Bacteriol 2000; 182: 6308– 6321
5. Kierek K, Watnick PI. The Vibrio cholerae O139 O-antigen polysaccharide is essential for Ca^{2+}-dependent biofilm development in sea water. Proc Natl Acad Sci USA 2003; 100: 14357– 14362
6. Hölzer SU, Schlumberger MC, Jäckel D, Hensel M. Effect of the O-antigen length of lipopolysaccharide on the functions of type III secretion systems in Salmonella enterica. Infect Immun 2009; 77: 5458– 5470
7. Morgenstein RM, Clemmer KM, Rather PN. Loss of the waaL O-antigen ligase prevents surface activation of the flagellar gene cascade in Proteus mirabilis. J Bacteriol 2010; 192: 3213– 3221
8. Post DM, Yu L, Krasity BC, Choudhury B, Mandel MJ et al. O-antigen and core carbohydrate of Vibrio fischeri lipopolysaccharide: composition and analysis of their role in Euprymna scolopes light organ colonization. J Biol Chem 2012; 287: 8515– 8530
9. Noel KD, Vandenbosch KA, Kulpaca B. Mutations in Rhizobium phaseoli that lead to arrested development of infection threads. J Bacteriol 1986; 168: 1392– 1401
10. Carlson RW, Kalembasa S, Turowski D, Pachori P, Noel KD. Characterization of the lipopolysaccharide from a Rhizobium phaseoli mutant that is defective in infection thread development. J Bacteriol 1987; 169: 4923– 4928
11. Cava JR, Elias PM, Turowski DA, Noel KD. Rhizobium leguminosarum CFN42 genetic regions encoding lipopolysaccharide structures essential for complete nodule development on bean plants. J Bacteriol 1989; 171: 8– 15 Forsberg LS, Bhat UR, Carlson RW. 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 2000; 275: 18851– 18863

12. Ojeda KJ, Simonds L, Noel KD. Roles of predicted glycosyltransferases in the biosynthesis of the Rhizobium etli CE3 O antigen. J Bacteriol 2013; 195: 1949– 1958 Valvano MA. Export of O-specific lipopolysaccharide. Front Biosci 2003; 8: s452– s471 Price NP, Momany FA. Modeling bacterial UDP-HexNAc: polyprenol-P HexNAc-1-P transferases. Glycobiology 2005; 15: 29R– 42R
13. Al-Dabbagh B, Mengin-Lecreulx D, Bouhss A. Purification and characterization of the bacterial UDP-GlcNAc:undecaprenyl-phosphate GlcNAc-1-phosphate transferase WecA. J Bacteriol 2008; 190: 7141– 7146
14. Merino S, Jimenez N, Molero R, Bouamama L, Regué M et al. A UDP-HexNAc:polyprenol-P GalNAc-1-P transferase (WecP) representing a new subgroup of the enzyme family. J Bacteriol 2011; 193: 1943– 1952
15. Patel KB, Ciepichal E, Swiezewska E, Valvano MA. 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 2012; 22: 116– 122
16. Li T, Simonds L, Kovrigin EL, Noel KD. In vitro biosynthesis and chemical identification of UDP-N-acetyl-d-quinovosamine (UDP-d-QuiNAc). J Biol Chem 2014; 289: 18110– 18120
17. Hwang S, Aronov A, Bar-Peled M. The Biosynthesis of UDP-d-QuiNAc in Bacillus cereus ATCC 14579. PLoS One 2015; 10: e0133790
18. Creuzenet C, Schur MJ, Li J, Wakarchuk WW, Lam JS. FlaA1, a new bifunctional UDP-GlcNAc C6 Dehydratase/ C4 reductase from Helicobacter pylori. J Biol Chem 2000; 275: 34873– 34880
19. Creuzenet C, Lam JS. Topological and functional characterization of WbpM, an inner membrane UDP-GlcNAc C6 dehydratase essential for lipopolysaccharide biosynthesis in Pseudomonas aeruginosa. Mol Microbiol 2001; 41: 1295– 1310
20. Olivier NB, Chen MM, Behr JR, Imperiali B. In vitro biosynthesis of UDP-N,N'-diacetylbacillosamine by enzymes of the Campylobacter jejuni general protein glycosylation system. Biochemistry 2006; 45: 13659– 13669 Schoenhofen IC, Mcnally DJ, Vinogradov E, Whitfield D, Young NM et al. Functional characterization of dehydratase/aminotransferase pairs from Helicobacter and Campylobacter: enzymes distinguishing the pseudaminic acid and bacillosamine biosynthetic pathways. J Biol Chem 2006; 281: 723– 732
21. Pinta E, Duda KA, Hanuszkiewicz A, Kaczyński Z, Lindner B et al. Identification and role of a 6-deoxy-4-keto-hexosamine in the lipopolysaccharide outer core of Yersinia enterocolitica serotype O:3. Chemistry 2009; 15: 9747– 9754
22. Noel KD, Forsberg LS, Carlson RW. Varying the abundance of O antigen in Rhizobium etli and its effect on symbiosis with Phaseolus vulgaris. J Bacteriol 2000; 182: 5317– 5324
23. Forsberg LS, Noel KD, Box J, Carlson RW. 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 2003; 278: 51347– 51359
24. Hartley MD, Larkin A, Imperiali B. Chemoenzymatic synthesis of polyprenyl phosphates. Bioorg Med Chem 2008; 16: 5149– 5156
25. Osborn MJ, Cynkin MA, Gilbert JM, Müller L, Singh M et al. Synthesis of bacterial O-antigens. In Methods in enzymology Academic Press; 1972; pp. 583– 601
26. Schäffer C, Wugeditsch T, Messner P, Whitfield C. Functional expression of enterobacterial O-polysaccharide biosynthesis enzymes in Bacillus subtilis. Appl Environ Microbiol 2002; 68: 4722– 4730
27. Kanjilal-Kolar S, Basu SS, Kanipes MI, Guan Z, Garrett TA et al. Expression cloning of three Rhizobium leguminosarum lipopolysaccharide core galacturonosyltransferases. J Biol Chem 2006; 281: 12865– 12878
28. Glover KJ, Weerapana E, Chen MM, Imperiali B. 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 2006; 45: 5343– 5350

29. Lukose V, Luo L, Kozakov D, Vajda S, Allen KN et al. Conservation and covariance in small bacterial phosphoglycosyltransferases identify the functional catalytic core. Biochemistry 2015; 54: 7326– 7334
30. Rush JS, Alaimo C, Robbiani R, Wacker M, Waechter CJ. A novel epimerase that converts GlcNAc-P-P-undecaprenol to GalNAc-P-P-undecaprenol in Escherichia coli O157. J Biol Chem 2010; 285: 1671– 1680
31. Cunneen MM, Liu B, Wang L, Reeves PR. Biosynthesis of UDP-GlcNAc, UndPP-GlcNAc and UDP-GlcNAcA involves three easily distinguished 4-epimerase enzymes, Gne, Gnu and GnaB. PLoS One 2013; 8: e67646
32. Vinogradov E, Frirdich E, Maclean LL, Perry MB, Petersen BO et al. Structures of lipopolysaccharides from Klebsiella pneumoniae. Eluicidation of the structure of the linkage region between core and polysaccharide O chain and identification of the residues at the non-reducing termini of the O chains. J Biol Chem 2002; 277: 25070– 25081
33. Casillo A, Parrilli E, Sannino F, Mitchell DE, Gibson MI et al. Structure-activity relationship of the exopolysaccharide from a psychrophilic bacterium: a strategy for cryoprotection. Carbohydr Polym 2017; 156: 364– 371
34. Noel KD, Sanchez A, Fernandez L, Leemans J, Cevallos MA. Rhizobium phaseoli symbiotic mutants with transposon Tn5 insertions. J Bacteriol 1984; 158: 148– 155
35. Lis M, Kuramitsu HK. The stress-responsive dgk gene from Streptococcus mutans encodes a putative undecaprenol kinase activity. Infect Immun 2003; 71: 1938– 1943
36. Dombrecht B, Vanderleyden J, Michiels J. Stable RK2-derived cloning vectors for the analysis of gene expression and gene function in gram-negative bacteria. Mol Plant Microbe Interact 2001; 14: 426– 430
37. Glazebrook J, Walker GC. Genetic techniques in Rhizobium meliloti. Methods Enzymol 1991; 204: 398– 418
38. Finan TM, Kunkel B, de Vos GF, Signer ER. Second symbiotic megaplasmid in Rhizobium meliloti carrying exopolysaccharide and thiamine synthesis genes. J Bacteriol 1986; 167: 66– 72
39. Patel KB, Valvano MA. In vitro UDP-sugar:undecaprenyl-phosphate sugar-1-phosphate transferase assay and product detection by thin layer chromatography. Methods Mol Biol 2013; 1022: 173– 183
40. El Ghachi M, Bouhss A, Blanot D, Mengin-Lecreulx D. The bacA gene of Escherichia coli encodes an undecaprenyl pyrophosphate phosphatase activity. J Biol Chem 2004; 279: 30106– 30113

SUPPLEMENTARY MATERIALS

Table S1. Bacterial strains containing homologs of R. etli CE3 WreQ, WreU and WreV.

WreQ (YP_470339.1), WreU (YP_471772.1), WreV (YP_471773.1) were used individually as the query sequences in BLAST searches against the non-redundant protein sequence (nr) database using the blastp algorithm. In the first WreQ and WreU BLAST searches, Rhizobiaceae were excluded (to eliminate the very high number of Rhizobium and closely affiliated rhizobial strains) and the max target sequences were 1000 and 5000 respectively. All saved strains from that search had WreU homologs with E-values of less than 1e-30; the WreQ homologs of these strains had E-values less than 1e-89. In the WreV BLAST search, initially the strategy was the same as with WreQ, and all the resulting sequences had an E-value of 0.0. However, the majority of sequences belonged to Pseudomonadaceae. Thus, a separate blast was performed to exclude Pseudomonadaceae as well, and with 5000 max target sequences, resulting in E-values less than 2e-127. Finally a search of only “Alpha-Proteobacteriaceae” was conducted and species, rather than strains, with homologs to all three genes were saved. In this final table, strains containing the three gene homologs are listed, with the accession numbers of the protein homologs. The E-values are less than 1e-50, 1e-30, and 4e-70 for matches with WreQ, WreU, and WreV, respectively. Where a strain number is not provided, it can be found through the gene accession numbers. The Phylo column indicates the phylogenetic affiliation of the strain within the domain Bacteria: FCB refers to the super-phylum FCB within the Bacteria; Alpha, Beta, Gamma, Delta, and Zeta refer to the sub-phyla within the Proteobacteria.

Strain	WreQ accession	WreU accession	WreV accession	Phylo
Chlorobium phaeobacteroides	WP_015961134.1	WP_041467797.1	WP_015961150.1	FCB
Pelodictyon phaeoclathratiforme	WP_012507375.1	WP_012507376.1	WP_012507339.1	FCB

Agrobacterium rhizogenes	WP_034519680.1	WP_034519733.1	WP_034514206.1	Alpha
Aureimonas altamirensis	WP_060606375.1	WP_060601527.1	WP_039195620.1	Alpha
Aureimonas frigidaquae	WP_062227553.1	WP_062227554.1	WP_062227987.1	Alpha
Azospirillum brasilense	WP_035680550.1	WP_014198799.1	WP_014240519.1	Alpha
Bradyrhizobium	WP_027564699.1	WP_028141810.1	WP_027564879.1	Alpha
Bradyrhizobium canariense	WP_085348811.1	WP_085384688.1	WP_085394538.1	Alpha
Bradyrhizobium erythrophlei	WP_079604300.1	WP_074274213.1	WP_092115479.1	Alpha
Candidatus Pelagibacter ubique	WP_029454191.1	WP_075521062.1	WP_029454193.1	Alpha
Mesorhizobium alhagi	WP_085984070.1	WP_008835798.1	WP_008835795.1	Alpha
Methylobacterium extorquens	WP_076643959.1	WP_015822390.1	WP_076643956.1	Alpha
Pseudooceanicola nitratireducens	WP_093454891.1	WP_093454896.1	WP_093447298.1	Alpha
Rhizobiales bacterium 63-7	OJU66100.1	OJU71699.1	OJU71698.1	Alpha
Rhizobium acidisoli	WP_054182687.1	WP_054185004.1	WP_054185003.1	Alpha
Rhizobium aegyptiacum	WP_064694690.1	WP_064696812.1	WP_064696703.1	Alpha
Rhizobium aethiopicum	WP_092748744.1	SCB62257.1	SCB58533.1	Alpha
Rhizobium alamii	WP_037099615.1	WP_051942476.1	WP_051942478.1	Alpha
Rhizobium bangladeshense	WP_064705588.1	WP_064712006.1	WP_064710004.1	Alpha
Rhizobium etli	WP_011426089.1	WP_011427467.1	WP_011427468.1	Alpha
Rhizobium favelukesii	WP_024317389.1	WP_024314414.1	WP_024317388.1	Alpha
Rhizobium freirei	WP_004127129.1	WP_037155461.1	WP_004126244.1	Alpha
Rhizobium gallicum	WP_018446548.1	WP_026230627.1	WP_083636055.1	Alpha
Rhizobium hainanense	WP_075853437.1	SCB21882.1	WP_075856697.1	Alpha
Rhizobium laguerreae	WP_077978790.1	WP_077979743.1	WP_077979742.1	Alpha
Rhizobium leguminosarum	WP_037140272.1	WP_027687178.1	WP_027687179.1	Alpha
Rhizobium marinum	WP_029621142.1	WP_029618564.1	WP_081852123.1	Alpha
Rhizobium mesoamericanum	WP_040678303.1	WP_007537568.1	WP_028745825.1	Alpha
Rhizobium miluonense	WP_092851701.1	SCB34436.1	WP_092843462.1	Alpha
Rhizobium multihospitium	WP_092720446.1	WP_092720448.1	WP_092714704.1	Alpha
Rhizobium phaseoli	WP_064818631.1	WP_085744233.1	WP_064837511.1	Alpha
Rhizobium tibeticum	WP_072375736.1	WP_072373828.1	WP_072375735.1	Alpha
Rhizobium tropici	WP_015341683.1	WP_015341684.1	WP_015338445.1	Alpha
Rhizobium/Agrobacterium group	WP_007694703.1	WP_007694702.1	WP_047524929.1	Alpha
Rhodopseudomonas palustris	WP_041798076.1	WP_080506310.1	WP_080674905.1	Alpha
Sphingopyxis macrogoltabida	WP_054589353.1	WP_054590413.1	WP_082396031.1	Alpha
Acidovorax soli	SEA14343.1	SEA14367.1	SEA14422.1	Beta
Acidovorax wautersii	WP_092939072.1	WP_092939071.1	WP_092939238.1	Beta
Burkholderia sp. 9120	WP_035552665.1	WP_035552664.1	WP_035552662.1	Beta
Burkholderiales bacterium RIFCSPLOWO2_12_FULL_61_40	OGB28531.1	OGB28530.1	OGB28529.1	Beta
Candidatus Accumulibacter sp. 66-26	OJW49362.1	OJW49363.1	OJW49596.1	Beta
Candidatus Methylopumilus turicensis	WP_082048396.1	WP_045751174.1	WP_045751173.1	Beta
Comamonadaceae bacterium CG1_02_60_18	OIN91272.1	OIN91271.1	OIN91269.1	Beta
Comamonadaceae bacterium CG2_30_57_122	OIP13220.1	OIP13227.1	OIP13228.1	Beta
Cupriavidus sp. WS	WP_081651598.1	WP_029044350.1	WP_043367598.1	Beta
Deefgea rivuli	WP_027469006.1	WP_027469005.1	WP_027469004.1	Beta
Herbaspirillum aquaticum	WP_088756155.1	WP_088753704.1	WP_088756171.1	Beta
Herbaspirillum frisingense	WP_039875505.1	WP_079216977.1	WP_079216979.1	Beta

Herbaspirillum rubrisubalbicans	WP_058894123.1	WP_058894118.1	WP_082686138.1	Beta
Herbaspirillum seropedicae	WP_069375417.1	WP_069375029.1	WP_069375030.1	Beta
Herbaspirillum sp. B39	WP_034337931.1	WP_034330935.1	WP_034330937.1	Beta
Herbaspirillum sp. HZ10	WP_088752243.1	WP_088752242.1	WP_088752241.1	Beta
Herbaspirillum sp. RV1423	WP_034299669.1	WP_081768878.1	WP_081768877.1	Beta
Herbaspirillum sp. WT00C	WP_075259658.1	WP_075255625.1	WP_075255626.1	Beta
Herbaspirillum sp. YR522	WP_008117337.1	WP_008113559.1	WP_008117346.1	Beta
Herminiimonas arsenicoxydans	WP_011870663.1	WP_011870664.1	WP_011870665.1	Beta
Herminiimonas arsenitoxidans	WP_076593873.1	WP_076590916.1	WP_076590917.1	Beta
Hydrogenophaga palleronii	WP_066269231.1	WP_066269232.1	WP_066269233.1	Beta
Hydrogenophaga sp. IBVHS1	WP_086126805.1	WP_086126802.1	WP_086126799.1	Beta
Hydrogenophaga sp. PML113	WP_083293270.1	WP_070400211.1	WP_070400209.1	Beta
Hydrogenophaga sp. RAC07	WP_083240110.1	WP_069049138.1	WP_069048529.1	Beta
Hydrogenophaga sp. Root209	WP_082585539.1	WP_056267373.1	WP_056267049.1	Beta
Hydrogenophaga sp. SCN 70-13	ODT34034.1	ODT34033.1	ODT34031.1	Beta
Janthinobacterium	WP_086138461.1	WP_086138462.1	WP_086138463.1	Beta
Janthinobacterium agaricidamnosum	WP_038489190.1	WP_038498699.1	WP_038489192.1	Beta
Janthinobacterium lividum	WP_070254296.1	WP_070254295.1	WP_070254294.1	Beta
Janthinobacterium sp. Ant5-2-1	WP_058049808.1	WP_058049809.1	WP_058049810.1	Beta
Janthinobacterium sp. GW458P	WP_086143996.1	WP_086143912.1	WP_086143913.1	Beta
Janthinobacterium sp. KBS0711	WP_046686130.1	WP_046686131.1	WP_046686132.1	Beta
Janthinobacterium sp. Marseille	WP_012080109.1	ABR88537.1	WP_012080107.1	Beta
Janthinobacterium sp. S3-2	WP_065307061.1	WP_065307060.1	WP_065307077.1	Beta
Janthinobacterium sp. TND4EL3	WP_076565808.1	WP_076565810.1	WP_076565792.1	Beta
Laribacter hongkongensis	WP_027824056.1	WP_027824055.1	WP_027824054.1	Beta
Laribacter hongkongensis	WP_088861422.1	WP_088861421.1	WP_088861420.1	Beta
Limnohabitans planktonicus	WP_053170360.1	WP_053170358.1	WP_053171553.1	Beta
Methylotenera mobilis	WP_041928542.1	WP_015831982.1	WP_015831983.1	Beta
Methylotenera sp. G11	WP_081987064.1	WP_047549180.1	WP_047549183.1	Beta
Methylotenera sp. L2L1	WP_036303317.1	WP_036303315.1	WP_036303313.1	Beta
Methylotenera sp. RIFCSPLOWO2_02_FULL_45_14	OGV78172.1	OGV78179.1	OGV77391.1	Beta
Methylotenera versatilis	WP_013147951.1	WP_041359898.1	WP_013147953.1	Beta
Methylovorus sp. MP688	WP_013440922.1	WP_013440921.1	WP_013442499.1	Beta
Noviherbaspirillum autotrophicum	WP_040038466.1	WP_040038467.1	WP_040038468.1	Beta
Noviherbaspirillum massiliense	WP_019141885.1	WP_026075971.1	WP_019141883.1	Beta
Pandoraea pnomenusa	WP_023598438.1	WP_023598439.1	WP_052167323.1	Beta
Pandoraea vervacti	WP_044458602.1	WP_044457406.1	WP_044458601.1	Beta
Paraburkholderia phenazinium	WP_090690488.1	WP_090690491.1	WP_090690498.1	Beta
Paraburkholderia sartisoli	WP_090532012.1	WP_090532010.1	WP_090532007.1	Beta
Paraburkholderia terrae	WP_086909626.1	WP_086909596.1	WP_086909597.1	Beta
Polaromonas glacialis	WP_029526176.1	WP_029526175.1	WP_029526173.1	Beta
Polaromonas jejuensis	WP_068831299.1	WP_068831300.1	WP_084389288.1	Beta
Polaromonas naphthalenivorans	WP_011802850.1	WP_011802849.1	WP_011802848.1	Beta
Polaromonas sp. A23	WP_077592725.1	WP_077592843.1	WP_077592727.1	Beta
Polaromonas sp. C04	WP_077563849.1	WP_077563851.1	WP_077563857.1	Beta
Polaromonas sp. CF318	WP_007862413.1	EJL78603.1	EJL78604.1	Beta
Polaromonas sp. CG9_12	WP_036808892.1	WP_036813202.1	WP_036808899.1	Beta
Polaromonas sp. EUR3 1.2.1	WP_026781163.1	WP_026781159.1	WP_026781158.1	Beta

Polaromonas sp. JS666	WP_011484870.1	WP_011484869.1	WP_011484867.1	Beta
Polaromonas sp. OV174	WP_091998443.1	WP_091998447.1	WP_091998450.1	Beta
Pseudorhodoferax sp. Leaf265	WP_056672590.1	WP_056672588.1	WP_056672598.1	Beta
Rhodoferax antarcticus	WP_075585179.1	WP_075585216.1	WP_075585168.1	Beta
Rhodoferax fermentans	WP_078364450.1	WP_078366834.1	WP_078364448.1	Beta
Rhodoferax ferrireducens	OQW88136.1	OQW86158.1	OQW88139.1	Beta
Rhodoferax sp. DCY110	WP_076197995.1	WP_076197993.1	WP_076198007.1	Beta
Sulfuriferula sp. AH1	WP_087447710.1	WP_087447709.1	WP_087447732.1	Beta
Thauera humireducens	WP_048706753.1	WP_048706755.1	AMO37809.1	Beta
Thauera terpenica	WP_040831089.1	WP_021250246.1	WP_021250247.1	Beta
Variovorax paradoxus	WP_057595238.1	WP_013543050.1	WP_062361990.1	Beta
Variovorax sp. CF313	WP_042673666.1	WP_007835336.1	WP_007835345.1	Beta
Variovorax sp. JS1663	WP_086924577.1	WP_086924597.1	WP_086924576.1	Beta
delta proteobacterium MLMS-1	WP_007292162.1	WP_007294049.1	WP_007294051.1	Delta
Geoalkalibacter ferrihydriticus	WP_074669617.1	WP_040101187.1	WP_052446538.1	Delta
Geobacter bemidjiensis	WP_012530059.1	WP_012530060.1	WP_012530067.1	Delta
Geobacter sp. DSM 9736	WP_088533764.1	WP_088536670.1	WP_088533767.1	Delta
Geobacter sp. M21	WP_015837859.1	WP_015837858.1	WP_015837853.1	Delta
Geobacteraceae bacterium GWB2_52_12	OGT95934.1	OGT95928.1	OGT95936.1	Delta
Alishewanella aestuarii	WP_008608033.1	WP_008608034.1	WP_008608039.1	Gamma
Alishewanella sp. HH-ZS	WP_065956080.1	WP_065956079.1	WP_065956072.1	Gamma
Arsukibacterium ikkense	WP_046556227.1	WP_046556226.1	WP_052748930.1	Gamma
Arsukibacterium sp. MJ3	WP_046555083.1	WP_046555081.1	WP_052750195.1	Gamma
Catenovulum maritimum	WP_048693212.1	WP_048693214.1	WP_048693216.1	Gamma
Chania multitudinisentens	WP_024911805.1	WP_024911804.1	WP_024911803.1	Gamma
Colwellia psychrerythraea	WP_033083676.1	WP_033093120.1	WP_033093119.1	Gamma
Colwellia sp. TT2012	WP_057830224.1	WP_057832342.1	WP_057830226.1	Gamma
Crenothrix polyspora	WP_087147254.1	WP_087147253.1	WP_087147252.1	Gamma
Cycloclasticus sp. DSM 27168	WP_073023169.1	WP_073023168.1	WP_073023165.1	Gamma
Endozoicomonas atrinae	WP_066015304.1	WP_066015305.1	WP_066015306.1	Gamma
Halomonas pantelleriensis	WP_089656468.1	WP_089656469.1	WP_089656471.1	Gamma
Halomonas salina	WP_040186301.1	WP_081945756.1	WP_040186321.1	Gamma
Lonsdalea quercina	WP_094107939.1	WP_094107961.1	WP_094107940.1	Gamma
Marinobacter hydrocarbonoclasticus	WP_011786059.1	WP_011786058.1	WP_011786057.1	Gamma
Marinobacter lipolyticus	WP_018405035.1	WP_018405034.1	WP_018405033.1	Gamma
Marinobacter salinus	WP_070965578.1	WP_070973522.1	WP_070965581.1	Gamma
Marinobacterium lutimaris	SEG57113.1	SEG57098.1	SEG57078.1	Gamma
Methyloglobulus morosus	WP_023495340.1	WP_023495341.1	WP_023495342.1	Gamma
Methylophaga sp. 41_12_T18	OUR72752.1	OUR72751.1	OUR72772.1	Gamma
Nitrincola nitratireducens	WP_036513279.1	EXJ09743.1	EXJ09744.1	Gamma
Oceanimonas sp. GK1	WP_041543213.1	WP_014293024.1	WP_014293023.1	Gamma
Oleiphilus	WP_068474530.1	WP_068553922.1	WP_068474520.1	Gamma
Photobacterium ganghwense	WP_047883801.1	WP_047883800.1	WP_047883802.1	Gamma
Photobacterium iliopiscarium	WP_045038329.1	WP_045038330.1	WP_045038331.1	Gamma
Photobacterium kishitanii	WP_065173147.1	WP_065173146.1	WP_065173145.1	Gamma
Photobacterium leiognathi	WP_023931563.1	WP_023931565.1	WP_045063689.1	Gamma
Photobacterium marinum	WP_007466443.1	WP_007466442.1	WP_036800127.1	Gamma
Photobacterium sp. SKA34	WP_006646475.1	WP_006646476.1	WP_039860628.1	Gamma

Photorhabdus asymbiotica	WP_065821825.1	WP_065821916.1	WP_065821915.1	Gamma
Photorhabdus asymbiotica	WP_036770100.1	WP_036770139.1	WP_036770093.1	Gamma
Pseudoalteromonas	WP_024589810.1	WP_024589809.1	WP_024589808.1	Gamma
Pseudoalteromonas lipolytica	WP_074989256.1	WP_074989257.1	WP_074989258.1	Gamma
Pseudoalteromonas luteoviolacea	WP_063378454.1	WP_063378455.1	WP_063378456.1	Gamma
Pseudoalteromonas luteoviolacea	WP_039608254.1	WP_039608255.1	WP_039608256.1	Gamma
Pseudoalteromonas rubra	WP_049864426.1	WP_046007017.1	WP_049864427.1	Gamma
Pseudoalteromonas sp. BMB	WP_069020842.1	WP_069020844.1	WP_069020848.1	Gamma
Pseudoalteromonas sp. BSi20429	WP_007584995.1	WP_007584994.1	WP_033026984.1	Gamma
Pseudomonas aeruginosa	WP_023130787.1	WP_023120375.1	WP_023130786.1	Gamma
Pseudomonas alcaliphila	WP_075747014.1	WP_064494598.1	WP_064494599.1	Gamma
Pseudomonas antarctica	WP_064451500.1	WP_064451502.1	WP_083359481.1	Gamma
Pseudomonas brassicacearum	ALQ02335.1	WP_025212585.1	WP_081352378.1	Gamma
Pseudomonas chengduensis	WP_090336757.1	WP_090336759.1	WP_090336760.1	Gamma
Pseudomonas chlororaphis	WP_053269669.1	WP_053280338.1	WP_053280337.1	Gamma
Pseudomonas composti	WP_074940188.1	WP_074940192.1	WP_074940422.1	Gamma
Pseudomonas fluorescens	WP_034155277.1	WP_034126702.1	WP_034126705.1	Gamma
Pseudomonas fluorescens	WP_003183858.1	WP_003189797.1	WP_003189805.1	Gamma
Pseudomonas frederiksbergensis	WP_076031558.1	WP_071554248.1	WP_071554254.1	Gamma
Pseudomonas oleovorans	WP_037054107.1	WP_037054105.1	WP_037054102.1	Gamma
Pseudomonas psychrotolerans	WP_058761729.1	WP_058760362.1	WP_058761296.1	Gamma
Pseudomonas putida	WP_069942133.1	WP_069942135.1	WP_069942131.1	Gamma
Pseudomonas sp. 1-7	KFJ93089.1	KFJ93088.1	KFJ93087.1	Gamma
Pseudomonas sp. C5pp	WP_039615100.1	WP_039615098.1	WP_039615101.1	Gamma
Pseudomonas sp. GM48	WP_007988308.1	WP_007985416.1	WP_007985421.1	Gamma
Pseudomonas sp. M30-35	WP_087516444.1	WP_087516445.1	WP_087516446.1	Gamma
Pseudomonas sp. MT-1	WP_045428458.1	WP_045428457.1	WP_045428456.1	Gamma
Pseudomonas sp. NFACC19-2	WP_072424717.1	WP_072424718.1	WP_072424719.1	Gamma
Pseudomonas stutzeri	WP_063540504.1	WP_063540515.1	WP_063540503.1	Gamma
Pseudomonas synxantha	WP_057022239.1	WP_057022149.1	WP_057022240.1	Gamma
Pseudomonas toyotomiensis	WP_074914515.1	WP_059391714.1	WP_059391715.1	Gamma
Pseudomonas veronii	WP_046483495.1	WP_046384274.1	WP_046384261.1	Gamma
Psychromonas aquimarina	WP_028862360.1	WP_028862361.1	WP_028862362.1	Gamma
Psychromonas ossibalaenae	WP_019615916.1	WP_019615917.1	WP_019615919.1	Gamma
Rheinheimera perlucida	WP_019674728.1	WP_019674726.1	WP_019674701.1	Gamma
Sedimenticola thiotaurini	WP_046859767.1	WP_046859766.1	WP_046859765.1	Gamma
Shewanella algae	WP_071237806.1	WP_071477444.1	WP_071237804.1	Gamma
Thioalkalivibrio sp. AKL10	WP_026340938.1	WP_081624528.1	WP_026340942.1	Gamma
Thioalkalivibrio sp. AKL12	WP_018951797.1	WP_018951790.1	WP_026289538.1	Gamma
Thioalkalivibrio sp. AKL6	WP_018145916.1	WP_026182128.1	WP_018145952.1	Gamma
Thioalkalivibrio sp. AKL7	WP_081624430.1	WP_019611620.1	WP_019612086.1	Gamma
Thioalkalivibrio sp. AKL8	WP_026304840.1	WP_081623074.1	WP_026304844.1	Gamma
Thioalkalivibrio sp. ALJ16	WP_018873978.1	WP_026280011.1	WP_018872892.1	Gamma
Thioalkalivibrio sp. ALM2T	WP_026333161.1	WP_026333164.1	WP_026333157.1	Gamma
Thioalkalivibrio sp. ALMg13-2	WP_019568703.1	WP_081620896.1	WP_026331226.1	Gamma
Thioalkalivibrio sp. ALMg2	WP_019563934.1	WP_081621424.1	WP_026330565.1	Gamma
Thioalkalivibrio sp. ALMg9	WP_018169554.1	WP_081620143.1	WP_026305952.1	Gamma
Thioalkalivibrio sp. K90mix	WP_012982927.1	WP_012982924.1	WP_012982951.1	Gamma
Thioalkalivibrio versutus	WP_047250985.1	WP_047250991.1	WP_047251909.1	Gamma

Thiothrix lacustris	OQX06987.1	OQX08772.1	OQX02620.1	Gamma
Vibrio alginolyticus	WP_086050204.1	WP_065646566.1	WP_065646570.1	Gamma
Vibrio anguillarum	AQP34955.1	WP_019282752.1	AQP34953.1	Gamma
Vibrio cholerae	WP_000433077.1	ADF80985.1	WP_076025069.1	Gamma
Vibrio crassostreae	WP_048661363.1	WP_048663200.1	WP_048663203.1	Gamma
Vibrio cyclitrophicus	WP_016791019.1	WP_016797283.1	WP_016797280.1	Gamma
Vibrio genomosp. F10	WP_017040691.1	WP_065576635.1	WP_065576634.1	Gamma
Vibrio harveyi CAIM1075	WP_050914487.1	WP_050937014.1	WP_050914489.1	Gamma
Vibrio metoecus	WP_055052054.1	WP_055044314.1	WP_055052055.1	Gamma
Vibrio mimicus	WP_000433076.1	WP_070382450.1	WP_000494954.1	Gamma
Vibrio nigripulchritudo	WP_022552005.1	WP_022552020.1	WP_022552006.1	Gamma
Vibrio parahaemolyticus	WP_020904316.1	WP_020904315.1	WP_020904314.1	Gamma
Vibrio splendidus	WP_017087969.1	WP_017085935.1	WP_017085937.1	Gamma
Vibrio tasmaniensis	WP_065106141.1	WP_041472952.1	WP_065113257.1	Gamma
Vibrio vulnificus	WP_013572513.1	WP_013572512.1	WP_013572511.1	Gamma
Mariprofundus ferrooxydans	WP_018294067.1	WP_026195384.1	WP_018294071.1	Zeta
Zetaproteobacteria bacterium CG2_30_46_52	OIP99107.1	OIP99141.1	OIQ00419.1	Zeta
Zetaproteobacteria bacterium TAG-1	WP_038246546.1	WP_038246540.1	WP_038249350.1	Zeta

Table S2. Bacterial strains containing homologs of R. etli CE3 WreQ, WreV and P. aeruginosa O6 WbpL. WreQ (YP_470339.1), WbpL (AAF23990.1) and WreV (YP_471773.1) were used individually as query sequences in BLAST searches against the non-redundant protein sequence (nr) database using the blastp algorithm. WreQ and WreV searches were performed following the same strategy as in Table S1. The WbpL BLAST search was performed by excluding Pseudomonadaceae and the maximum target sequences were set at 1000. The resulting matches to WbpL all have E-values less than 7e-32. Strains containing homologs to the three genes are listed below, along with the accession numbers of the protein homologs and the phylogenetic affiliations of the strains. Where strain numbers are absent, they can be found by investigating the gene accensions.

Strain	WreQ accession	WbpL accession	WreV accession	Phylo
Alphaproteobacteria bacterium	OFW83437.1	OFW83436.1	OFW84297.1	Alpha
Azospirillum humicireducens	WP_063635086.1	WP_063635083.1	WP_063635087.1	Alpha
Azospirillum lipoferum	WP_014248476.1	WP_012974782.1	WP_085556015.1	Alpha
Azospirillum oryzae	WP_085089998.1	WP_085089995.1	WP_085089999.1	Alpha
Azospirillum sp. B506	WP_042696170.1	WP_042696173.1	WP_042696169.1	Alpha
Azospirillum thiophilum	WP_045580795.1	WP_045580792.1	WP_045580796.1	Alpha
Devosia chinhatensis	WP_046103598.1	WP_046103597.1	WP_046103596.1	Alpha
Devosia chinhatensis	WP_046103598.1	WP_046103597.1	WP_046103596.1	Alpha
Devosia sp. 67-54	OJX19969.1	OJX19967.1	OJX19968.1	Alpha
Devosia sp. Root105	WP_082528546.1	WP_055884710.1	WP_055872943.1	Alpha
Haematospirillum jordaniae	WP_066132823.1	AMW34221.1	WP_066133023.1	Alpha
Magnetospirillum gryphiswaldense	WP_041634754.1	WP_041634753.1	WP_052589129.1	Alpha
Magnetospirillum gryphiswaldense	CDL00675.1	CDL00674.1	CDL00757.1	Alpha
Magnetospirillum sp. XM-1	WP_068430147.1	CUW37957.1	WP_068430462.1	Alpha
Mesorhizobium sp. YR577	WP_091914561.1	WP_091914110.1	WP_091914112.1	Alpha
Mesorhizobium sp. YR577	WP_091914561.1	WP_091914110.1	WP_091914112.1	Alpha
Rhizobium sp. Root149	WP_062557229.1	WP_062557228.1	WP_062557230.1	Alpha
Rhodocista sp. MIMtkB3	WP_075773469.1	WP_075772476.1	WP_075769167.1	Alpha
Rhodospirillaceae bacterium	OUX31398.1	OUX31399.1	OUX31397.1	Alpha

Rhodospirillales bacterium	OHC75555.1	OHC75553.1	OHC75556.1	Alpha
Terasakiella sp. PR1	WP_069186887.1	WP_069186886.1	WP_069186888.1	Alpha
Thalassospira alkalitolerans	WP_085616563.1	WP_085615659.1	WP_085614765.1	Alpha
Thalassospira mesophila	WP_085582449.1	WP_085578919.1	WP_085578921.1	Alpha
Thalassospira profundimaris	WP_064788430.1	WP_008888595.1	WP_008888594.1	Alpha
Thalassospira sp. MCCC	WP_085588244.1	WP_085590024.1	WP_085590155.1	Alpha
Thalassospira sp. Nap_22	KXJ56428.1	KXJ54524.1	KXJ54525.1	Alpha
Thalassospira sp. TSL5-1	WP_073955943.1	WP_073956208.1	WP_073955272.1	Alpha
Thalassospira xiamenensis	WP_062958433.1	WP_062960508.1	WP_062960509.1	Alpha
Azoarcus toluclasticus	WP_040394211.1	WP_051092234.1	WP_018988035.1	Beta
Lautropia sp. SCN 70-15	ODT33269.1	ODT33266.1	ODT33268.1	Beta
Methylibium sp. NZG	KNZ31778.1	KNZ31777.1	KNZ31776.1	Beta
Aeromonas veronii	WP_005350303.1	WP_005338066.1	WP_005338065.1	Gamma
Aquisalimonas asiatica	SEO80355.1	SEO80337.1	SEO80272.1	Gamma
C. Tenderia electrophaga	ALP53219.1	ALP54791.1	ALP53220.1	Gamma
C. Thiodiazotropha endoloripes	WP_068994608.1	WP_068994607.1	WP_068994639.1	Gamma
Cellvibrio sp. PSBB006	WP_087465803.1	WP_087465802.1	WP_087465797.1	Gamma
Crenothrix sp. D3	OTE95239.1	OTE95238.1	OTE95237.1	Gamma
Ectothiorhodospira marina	WP_090255654.1	WP_090255652.1	WP_090250421.1	Gamma
gamma proteobacterium IMCC2047	EGG98104.1	EGG99498.1	EGH00084.1	Gamma
Halomonas utahensis	WP_077530405.1	WP_077530383.1	WP_077530382.1	Gamma
Immundisolibacter cernigliae	WP_068803522.1	WP_068803523.1	WP_083214804.1	Gamma
Legionella geestiana	WP_028385720.1	WP_051550994.1	WP_051550925.1	Gamma
Legionella israelensis	WP_058500563.1	WP_058500543.1	WP_058500737.1	Gamma
Marinobacter mobilis	WP_091817059.1	WP_091817062.1	WP_091817065.1	Gamma
Methylobacter tundripaludum	WP_006890946.1	WP_006890945.1	WP_006890944.1	Gamma
Methylomicrobium	WP_005371725.1	WP_005371723.1	WP_005371722.1	Gamma
Methylomicrobium alcaliphilum	WP_014146586.1	WP_046061293.1	WP_046060910.1	Gamma
Methylomicrobium buryatense	WP_017841410.1	WP_040575735.1	WP_017841407.1	Gamma
Methylosarcina fibrata	WP_020562151.1	WP_020562152.1	WP_020562153.1	Gamma
Methylosarcina lacus	WP_029646324.1	WP_024296937.1	WP_024296936.1	Gamma
Methylovulum miyakonense	WP_019868468.1	WP_019868467.1	WP_019868466.1	Gamma
Methylovulum psychrotolerans	WP_088618707.1	WP_088618706.1	WP_088618705.1	Gamma
Microbulbifer variabilis	WP_020411331.1	WP_026304876.1	WP_020411328.1	Gamma
Motiliproteus sp. MSK22-1	WP_076720080.1	WP_076720081.1	WP_076720083.1	Gamma
Rheinheimera sp. SA_1	WP_082971652.1	WP_068063778.1	WP_082971650.1	Gamma
Stenotrophomonas rhizophila	WP_061203702.1	KWW13329.1	WP_061203704.1	Gamma
Thioalkalivibrio sp. AKL11	WP_083916429.1	WP_026288304.1	WP_018941009.1	Gamma
Thioalkalivibrio sp. AKL17	WP_026289054.1	WP_018947748.1	WP_018945921.1	Gamma
Thioalkalivibrio sp. ALJ15	WP_083925125.1	WP_020146293.1	WP_020146532.1	Gamma
Thioalkalivibrio sp. ALJ2	WP_081616998.1	WP_018994651.1	WP_018994650.1	Gamma
Thioalkalivibrio sp. ALR17-21	WP_024329946.1	WP_024329948.1	WP_024328922.1	Gamma
Thiohalomonas denitrificans	WP_092998985.1	WP_092999077.1	WP_092992255.1	Gamma
Deltaproteobacteria bacterium	OGQ52421.1	OGQ54831.1	OGQ57681.1	Delta

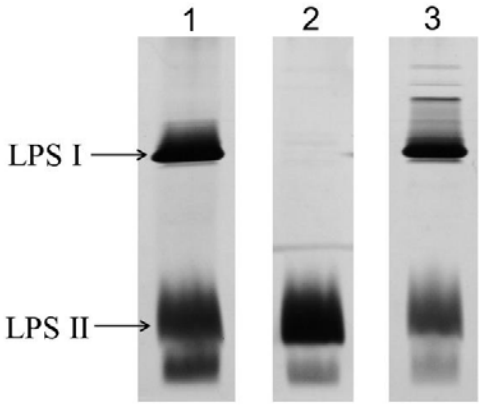


Fig. S1. Complementation of the *R. etli wreU*::Km mutant strain CE566 with *His$_6$-wreU*. Silver-stained SDS-PAGE (18% gel) of LPS samples from whole-cell lysates. Lanes: lane 1, CE3 (wild-type); lane 2, CE566 (*wreU*::Km); lane 3, CE566/pLS22 (*wreU*::Km/*His$_6$-wreU*). LPS I carries O antigen, whereas LPS II does not [1, 2].

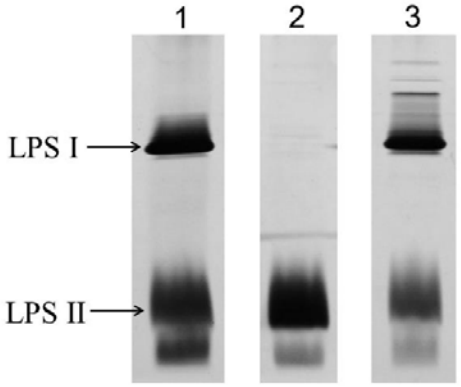
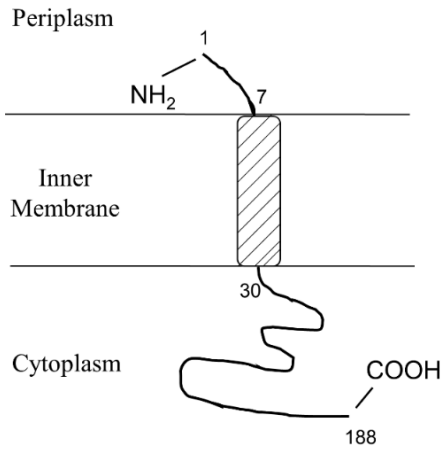


Fig. S2. Complementation of the *R. etli wreG*::Tn*5* mutant strain CE358 with *wreG-His$_6$*. Silver-stained SDS-PAGE (18% gel) of LPS samples from whole-cell lysates. Lanes: lane 1, CE3 (wild-type); lane 2, CE358 (*wreG*::Tn*5*); lane 3, CE358/pTL63 (*wreG*::Tn*5*/*wreG-His$_6$*).

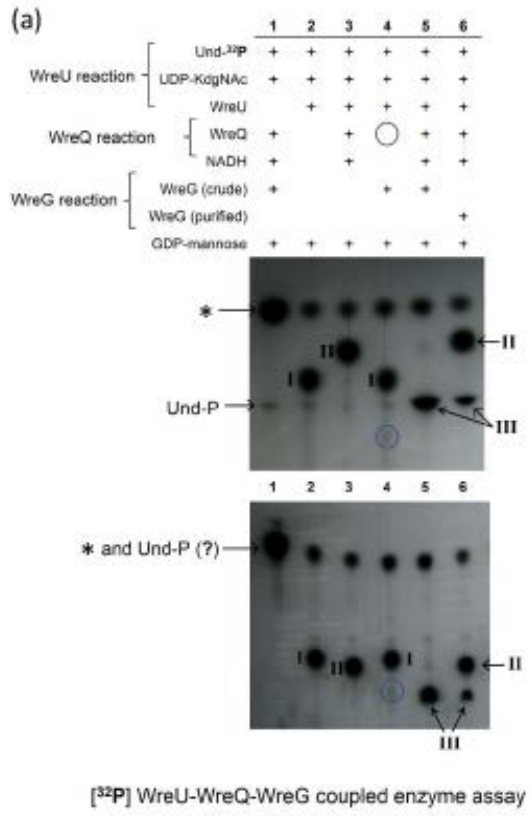
(a) WreU (188 aa)



(b)

```
WreU  ---MGLKRAVDFLLALIASVILLVPILVVALSVRLTSPGPILYWSKRIGR  47
WbaP  RSSRFLKRTFDIVCSIMILIIASPLMIYLWYKV-TRDGGPAIYGHQRVGR  322
PglC  MYEKVFKRIFDFILALVLLVLFSPVILITALLLKI-TQGSVIFTQNRPGL  49
WecP  PIYAFIKRGMDIIAAVIAIPLFSPLMIATAVLIKLESPGPVMFLQNRVGK  278

WreU  FNQIFLMPKFRSMRVDTPTVATHLL---------------ENPERFLTPI  82
WbaP  HGKLFPCYKFRSMVMNSQEVLKELLANDPIARAEWEKDFKLKNDPRITAV  372
PglC  DEKIFKIYKFRTMSDERDEKG---------------ELLS--DELRLKAF  82
WecP  GNRDFRIYKFRSMCQNSEQHG---------------AQFAQDGDMRVTRV  313

WreU  GSFLRKSSLDELPQLWCILAGKMSFVGPRPALYNQYDLIELRTVY--GVD  130
WbaP  GRFIRKTSLDELPQLFNVLKGDMSLVGPRPIVSDELERYCDDVDYYL---  419
PglC  GKIVRSLSLDELLQLFNVLKGDMSFVGPRPLLVEYLPLYNKEQ---KLRH  129
WecP  GKVIRKLRIDELPQFFNVLKGDMSLIGPRPEQRTFVDQFDREIPFYMYRH  363

WreU  KLLPGLTGWAQINGRDE---LPIPEKVKFDVEYLERRSLGFDMRILFLTA  177
WbaP  MAKPGMTGLWQVSGRND---VDYDTRVYFDSWYVKNWTLWNDIAILFKTA  466
PglC  KVRPGITGWAQVNGRNA---ISWQKKFELDVYYVKNISFLLDLKIMFLTA  176
WecP  IVRPGISGWAQVVHGYAADADDTRIKIEHDFYYIKNFSLWLDVLIVFKTI  413

WreU  EKVVRRKGIKH--------------                            188
WbaP  KVVLRRDGAY---------------                            476
PglC  LKVLKRSGVSKEGHVTTEKFNGKN                             200
WecP  RTILTGFGAR--------------                             423
```

Fig. S3. (a) Topological model for R. etli CE3 WreU. The shaded rectangle boxes represent transmembrane (TM) segments. The numbers indicate the amino acid positions of the boundary of each TM domain. (b) Multiple sequence alignment of the R. etli CE3 WreU (YP_471772) with C. jejuni PglC (YP_002344517), the C-terminal domain (274-476 aa) of S. enterica WbaP (NP_461027) and the C-terminal domain (229-423) of A. hydrophila WecP (EU274663), by use of the Clustal Omega program [3] (available on the World Wide Web at www.ebi.ac.uk/Tools/msa/clustalo/). Identical amino acids are highlighted in black and similar amino acids are highlighted in grey.



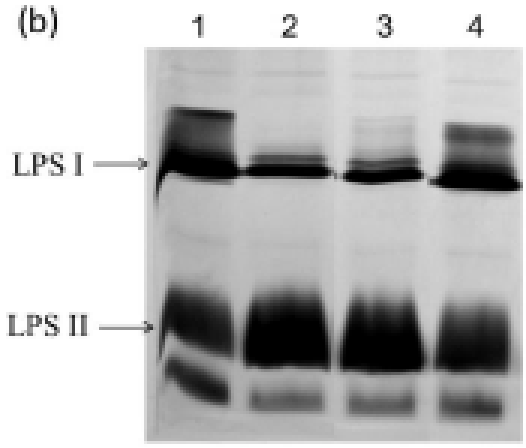


Fig. S4. (a) Highlight of the possible Und-PP-KdgNAc-Man product present in Fig. 5. (b) Suppression of R. etli wreQ::Tn5 mutant strain CE166 by introducing multiple copies of wreU or wreG gene. Lanes: lane 1, CE3 (R. etli wild-type strain); lane 2, CE166 (wreQ::Tn5); lane 3, CE166/pLS22, strain CE166 harboring pLS22 plasmids expressing His6-WreU; lane 4, CE166/pTL63, strain CE166 harboring pTL63 plasmids expressing WreG-His6. Note that extra copies of wreG suppressed the deficiency in O-antigen amount (lane 4), whereas extra copies of wreU did not (lane 3). Hypothesis 1 of Fig. 1(c) would have predicted suppression by multiple copies of wreU. A probable explanation of the suppression is that higher concentrations of WreG compensate for its weaker catalysis when KdgNAc is the acceptor residue in the reaction instead of the preferred QuiNAc residue.

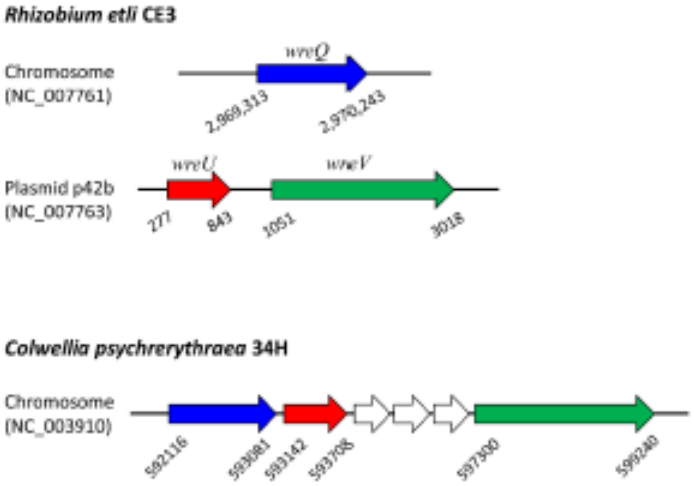


Fig. S5. The genomic locations of Rhizobium etli strain CE3 wreV, wreU, wreQ genes and their orthologs in Colwellia psychrerythraea strain 34H. wreV, wreU, wreQ and their orthologs are shown in green, red, blue respectively. Protein accession numbers: R. etli CE3 WreV (YP_471773), WreU (YP_471772), WreQ (YP_470339); C. psychrerythraea 34H WreV ortholog (WP_011041444), WreU ortholog (WP_011041440), WreQ ortholog (WP_011041439).

REFERENCES

1. Carlson RW, Kalembasa S, Turowski D, Pachori P, Noel KD. Characterization of the lipopolysaccharide from a Rhizobium phaseoli mutant that is defective in infection thread development. J Bacteriol 1987;169(11):4923-4928.
2. Noel KD, Forsberg LS, Carlson RW. Varying the abundance of O antigen in Rhizobium etli and its effect on symbiosis with Phaseolus vulgaris. J Bacteriol 2000;182(19):5317-5324.
3. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 2011;7:539.