NIH Public Access
Author Manuscript
Biochemistry. Author manuscript; available in PMC 2013 July 24.

Published in final edited form as:
Biochemistry. 2012 July 24; 51(29): 5804–5810. doi:10.1021/bi300640b.

The *wciN* gene encodes an α-1,3-galactosyltransferase involved in the biosynthesis of the capsule repeating unit of *Streptococcus pneumoniae* serotype

Weiqing Han[a,b,†], **Li Cai**[c,†], **Baolin Wu**[b], **Lei Li**[a], **Zhongying Xiao**[b], **Jiansong Cheng**[a,*], and **Peng G. Wang**[b,*]

[a]State Key Laboratory of Medicinal Chemical Biology, Nankai university, Tianjin 300071, China

[b]Department of Chemistry, Georgia State University, Atlanta, GA 30302, USA

[c]Department of Chemistry, University of South Carolina Salkehatchie, Walterboro, SC 29488, USA

Abstract

Almost all *Streptococcus pneumoniae* (pneumococcus) capsule serotypes employ the Wzy-dependent pathway for their capsular polysaccharide (CPS) biosynthesis. The assembly of the CPS repeating unit (RU) is the first committed step in this pathway. The *wciN* gene was predicted to encode a galactosyl transferase involved in the RU assembly of pneumococcus type 6B CPS. Herein, we provide the unambiguous *in vitro* biochemical evidence that *wciN* encodes an α-1,3-galactosyl transferase catalyzing the transfer of galactosyl from UDP-Gal onto the Glcα-pyrophosphate-lipid (Glcα-PP-lipid) acceptor to form Galα(1-3)Glcα-PP-lipid. A chemically synthesized acceptor (Glcα-PP-O$(CH_2)_{10}CH_3$) was used to characterize the WciN activity. The disaccharide product, i.e. Galα (1-3)Glcα-PP-O$(CH_2)_{10}CH_3$, was characterized by mass and NMR spectroscopy. Substrate specificity study indicated that the acceptor structural region composed of pyrophosphate and lipid moieties may play an important role in the enzyme-acceptor recognition. Furthermore, divalent metal cations were found indispensable to the WciN activity, suggesting that this glycosyltransferase (GT) belongs to the GT-A superfamily. By analyzing the activities of six WciN mutants, a DXD motif involved in the coordination of a divalent metal cation was identified. This work provides a chemical biology approach to characterize the activities of pneumococcal CPS GTs *in vitro*, and will help to understand the pneumococcal CPS biosynthetic pathway.

Streptococcus pneumoniae (pneumococcus), a leading bacterial pathogen worldwide, is coated by a thick layer of capsular polysaccharide (CPS) on the cell surface. The extracellular CPS is the major pneumococcal virulence factor and greatly contributes to the pneumococcal pathogenicity by enhancing the resistance to the complement-mediated opsonophagocytosis.[1,2] To date, 91 different CPS serotypes have been identified with each

[*]Corresponding author: (1) Jiansong Cheng, State Key Laboratory of Medicinal Chemical Biology, Nankai University, 94 Weijin Road, Nankai District, Tianjin 300071, P. R. China. Tel: +86 22 23507880; jiansongcheng@nankai.edu.cn. (2) Peng George Wang, Department of Chemistry, Georgia State University, P.O. Box 4098, Atlanta, Georgia 30302-40, USA. Tel.: +1 404 4133591; Fax: +1 404 4135505; Office: 787 Petit Science Center; pwang11@gsu.edu.
[†]W. H. and L. C. contributed equally to this work.

Supporting Information

The synthetic scheme (Scheme S1) and detailed synthesis procedures of Glcα-PP-O$(CH_2)_{10}CH_3$. The possible reason leading to the alkali-instability of Glcα-PP-O$(CH_2)_{10}CH_3$ (Fig. S1). Oligonucleotide primers used in this study (Table S1). NMR spectrometry data for various compounds including intermediates and products. This material is available free of charge via the Internet at http://pubs.acs.org.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

serotype expressing a serologically distinct polysaccharide.[3] Among all pneumococcal CPS serotypes, only two (types 3 and 37) are synthesized through the synthase-dependent pathway, while all others employ the Wzy-dependent pathway.[4,5] The Wzy-dependent pathway starts with the ligation of a sugar 1-phosphate onto the undecaprenyl phosphate (Und-P) lipid to form an Und-PP-sugar intermediate in the cytosolic surface of the cytoplasmic membrane, catalyzed by an integral membrane protein.[6,7] In pneumococcus, WchA fulfills this role by transferring Glc-1-P onto Und-P (Fig. 1).[8,9] Following this initiation step, specific glycosyltransferases (GTs) sequentially add different sugar residues onto the Und-PP-sugar intermediate, forming an oligosaccharide repeating unit (RU).[10–12] These RUs are then translocated by flippase Wzx to the outer surface of the cytoplasmic membrane,[13,14] where they become polymerized by polymerase Wzy (Fig. 1).[15,16] In pneumococcus, the mature CPS is then translocated by the Wzd/Wze complex to the cell surface and attached to peptidoglycan (Fig. 1).[5]

Currently widely used pneumococcal vaccines including polysaccharide vaccines and polysaccharide-protein conjugate vaccines are both derived from CPSs.[17] A better understanding of the pneumococcal CPS biosynthetic pathway may help to elaborate new anti-infectious agents, or even provide novel biotechnological tools to develop more efficient vaccines.[18,19] The investigations on GT functions represent an initial but significant step in the efforts to fully understand the pneumococcal CPS biosynthesis. Moreover, GTs have become powerful synthetic tools for the preparation of oligosaccharides or glycoconjugates.[20] The highly diverse activities of pneumococcal CPS GTs provide a large reservoir of GTs for such a purpose. However, the lack of knowledge of the pneumococcal GT activities has remarkably limited their applications. Taken together, more investigations on pneumococcal GT functions should be performed.

While the CPS genetic loci of all pneumococcal serotypes have been previously sequenced and gene functions also proposed,[4,5,21] only a limited number of pneumococcal GTs have been characterized *in vitro*.[2,22] Such limitations have primarily resulted from the barrier in obtaining appropriate acceptor substrates. The natural acceptor substrates of pneumococcal GTs contain a 55-carbon Und lipid (Fig. 2). Such substrates are difficult to handle owing to the extremely poor solubility in aqueous media. Herein, we chemically synthesized an acceptor with a shorter 11-carbon alkyl chain as an Und surrogate (Glcα-PP-O$(CH_2)_{10}CH_3$) to characterize a GT (i.e. WciN) involved in the pneumococcus type 6B CPS biosynthesis. Type 6B is one of the 23 serotypes that are most commonly associated with pneumococcal diseases,[17] and its CPS RU is as follow: Rhaα (1-4)Rib-ol(5-P-2)Galα (1-3)Glcα-PP-Und (Fig. 2).[23,24] Three GTs, namely WciN, WciO and WciP, were proposed to sequentially add three sugar residues onto Glcα-PP-Und to form the complete type 6B RU (Fig. 2).[21] However, none of them has been functionally confirmed. Using the synthetic acceptor, we biochemically identified WciN as a UDP-Gal:Glcα-PP-lipid α1,3-galactosyl transferase (Fig. 3). Because the RUs of most pneumococcal serotypes use Glc as the reducing-end residue,[21] this synthetic acceptor can be used as a general probe to characterize the activities of many other pneumococcal GTs.

Materials and Methods

Bacterial strains, plasmids, and materials

Pneumococcus type 6B strain was kindly provided by C. J. Orihuela from the University of Texas Health Science Center. *Escherichia coli* DH5α competent cells were from Gibco-BRL Life Technology. *E. coli* BL21(DE3) competent cells were from Novagen Inc. (Madison, WI). The Site-Directed Mutagenesis kit was purchased from Agilent Technologies. Expression plasmid pET-15b was purchased from Novagen (Carlsbad, CA).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

All kits and enzymes were used following manufacturers' instructions. All other chemical reagents and solvents were obtained from Sigma Aldrich unless otherwise noted.

Cloning, expression, and purification of WciN

The *wciN* gene was amplified by PCR from the genomic DNA of pneumococcus type 6B using the cloning primers listed in Table S1 of the Supporting Information. The PCR product was digested with *Nde*I and *BamH*I and subsequently inserted into the pET-15b vector linearized by the same restriction enzymes, forming a pET-15b-*wciN* recombinant plasmid. The recombinant plasmid was confirmed by restriction mapping and sequencing. The correct construct was subsequently transformed into *E. coli* BL21(DE3) for protein expression.

E. coli BL21(DE3) harboring the recombinant plasmid was incubated in 1 L of Luria–Bertani (LB) medium supplemented with 100 μg/mL ampicillin at 37 °C until A_{600} reached ~0.6. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was then added to a final concentration of 0.2 mM. The protein expression was allowed to continue for 20 h at 16 °C. Cells were harvested and washed with 20 mM Tris–HCl, pH 7.5. The cell pellet was resuspended in the buffer (50 mM Tris–HCl, pH 7.5, 0.3 M NaCl, 10 mM imidazole, 5% glycerol and 1 mM PMSF) and disrupted by sonication as previously described.[25] Cell lysates were centrifuged at 15,000 × g for 20 min at 4 °C to remove the cell debris. The resulting supernatant was applied onto a 5-mL HisTrap HP column (GE Healthcare) that was pre-equilibrated with the buffer containing 50 mM Tris–HCl, pH 7.5, 0.3 M NaCl, 5% glycerol and 10 mM imidazole. The column was washed with 10 column volumes of 50 mM imidazole in the same buffer, and the protein was eluted with the buffer containing 250 mM imidazole. Finally, eluted WciN was applied to the PD-10 column (GE Healthcare) to desalt, and then eluted with storage buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 20% glycerol). Purified protein was analyzed by SDS–PAGE and kept at −20 °C.

Chemical synthesis of the Glcα-PP-O$(CH_2)_{10}CH_3$ acceptor

The synthetic scheme (Scheme S1 of the Supporting Information), procedures and product characterization can be found in the Supporting Information. Briefly, peracetylated Glcα-PP-O$(CH_2)_{10}CH_3$ was obtained via the coupling of protected glucose 1-phosphate and undecanol 1-phosphate. Peracetylated glucose 1-phosphate was prepared via phosphoramide chemistry.[10] A final deacetylation step of coupling product afforded desired Glcα-PP-O$(CH_2)_{10}CH_3$. It is notable that, cation-exchange resin (pyridinium form, weak acidity) should be immediately added to neutralize the reaction mixture upon the completion of the Zemplen deacetylation of peracetylated Glcα-PP-O$(CH_2)_{10}CH_3$ to avoid product degradation (the Supporting Information).

Activity assay of WciN using the synthetic acceptor

The reaction was performed with a 100-μL mixture containing 50 mM Tris, pH 7.5, 10 mM $MgCl_2$, 5 mM UDP-Gal, 5 mM Glcα-PP-O$(CH_2)_{10}CH_3$, and ~20 ug purified WciN. The reaction was allowed to continue for 3 h at 37 °C and the product was detected using high resolution mass spectrometry (HRMS).

Preparative-scale synthesis of disaccharide product Galα (1,3)Glcα-PP-O$(CH_2)_{10}CH_3$

A 8-mL reaction mixture containing 50 mM Tris, pH 7.5, 10 mM 5 mM $MgCl_2$, UDP-Gal, 5 mM Glcα-PP-O$(CH_2)_{10}CH_3$, and ~3 mg purified WciN was incubated at 37 °C with mild agitation until substrates were observed completely converted to the product as monitored by mass spectroscopy. The reaction was then quenched by adding an equal volume of cold methanol and precipitated WciN was removed by brief centrifugation. The supernatant was

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

loaded on to a C18 reverse phase column (8 mm × 80 mm, particle size 40 mm, pore size 60 A, from Aldrich) and eluted with a gradient of CH_3OH (0–100% CH_3OH) in water containing 10 mM NH_4HCO_3 (the product was eluted out at ~50% CH_3OH). The fractions containing desired product were pooled and concentrated to give the product Galα (1,3)Glcα-PP-O$(CH_2)_{10}CH_3$ (16.5 mg, 60%). ^{1}H NMR (500 MHz, D_2O) δ 5.65 (dd, J = 7.2, 3.5 Hz, 1H), 5.47 (d, J = 3.9 Hz, 1H), 4.31 (t, J = 6.5 Hz, 1H), 4.06 (d, J = 3.2 Hz, 1H), 3.95–4.02 (m, 5H), 3.87–3.92 (m, 2H), 3.83 (dd, J = 12.3, 4.4 Hz, 1H), 3.77–3.79 (m, 2H), 3.74 (app t, J = 9.9 Hz, 1H), 3.67 (dt, J = 9.6, 3.2 Hz, 1H), 1.67–1.72 (m, 2H), 1.29–1.43 (m, 16H), 0.90 (t, J = 7.1 Hz, 3H); ^{13}C NMR (125 MHz, D_2O) δ 99.0, 95.7 (d), 78.9, 72.6, 70.6, 70.4 (d), 69.8, 69.4, 69.2, 68.7, 67.1 (d), 60.8, 60.2, 31.2, 29.8 (d), 28.8 (multiple carbons), 28.7, 28.5 (multiple carbons), 24.9, 22.1, 13.4; HRMS (ESI) calcd for $C_{23}H_{45}O_{17}P_2$ $(M\text{-}H)^-$ 655.2137, found 655.2125 *m*/*z*.

Site-directed mutagenesis of WciN

Six WciN mutants, $WciN_{D28A}$, $WciN_{D30A}$, $WciN_{D102A}$, $WciN_{D104A}$, $WciN_{D116A}$, and $WciN_{D118A}$, were created by using the QuikChange Site-Directed Mutagenesis kit (Agilent Technologies). Briefly, oligonucleotide primers (Table S1 of the Supporting Information) were designed to create the desired amino acid change. The pET-15b-*wciN* plasmid was used as the DNA template. PCR products were digested with *Dpn*I, and transformed into XL1-Blue Supercompetent Cells. Plasmids were isolated and sequenced to confirm the desired substitution of amino acid residues. The WciN mutants were expressed, purified, and applied to the *in vitro* WciN activity assay as described before.

Results

Expression and purification of WciN

The recombinant plasmid (pET-15b-*wciN*) encodes full-length WciN with an N terminus 6×His epitope. The expression of WciN at a low temperature (16 °C) significantly increased the amount of soluble portion of the overall protein. The His_6-WciN fusion protein could be purified to near homogeneity in one step using Ni-column with a yield of ~5 mg/L in the LB medium (Fig. 4). The recombinant protein has an apparent molecular weight of 39 kDa as estimated by SDS–PAGE (Fig. 4), close to the calculated value (38.799 kDa). The sequence analysis showed that WciN contains no transmembrane segments, consistent with the observation that WciN could be purified without the need of any detergents.

Characterization of the WciN activity using the synthetic acceptor substrate

The chemically synthesized acceptor (Glcα-PP-O$(CH_2)_{10}CH_3$) was used to characterize the WciN activity. The disaccharide product was first examined using high resolution mass spectrometry. A prominent peak with m/z ratio at 655.2125 [calcd for $C_{23}H_{45}NO_{11}P_2$ $(M\text{-}H)^-$ 655.2137] was observed, consistent with the formation of the Galα (1,3)Glcα-PP-O$(CH_2)_{10}CH_3$ product (Fig. 5). The structure of the product was further determined by NMR. A reasonable amount of Galα (1,3)Glcα-PP-O$(CH_2)_{10}CH_3$ product was obtained from preparative scale enzymatic synthesis. ^{1}H NMR, ^{13}C NMR, H-H COSY, HSQC, and HMBC spectra of Galα (1,3)Glcα-PP-O$(CH_2)_{10}CH_3$ were provided in the Supporting Information. For the anomeric proton of Gal (δ 5.47), a coupling constant of 3.9 Hz indicates an a-linkage (Fig. 6a). The Gal1→3Glc linkage was confirmed by HMBC experiment, where a strong coupling was clearly observed between the Gal C-1 and Glc H-3 signals (Fig. 6b).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Substrate specificity of WciN

The specificity of WciN towards different acceptor substrates was further investigated using compounds shown in Table 1. Out of 8 chosen acceptors, only Glcα-PP-O$(CH_2)_{10}CH_3$ was accepted by WciN, while all others do not show any detectable activity (Table 1). The absence of the WciN activity on entries 4–8 suggests that the structural region composed of pyrophosphate and lipid moieties may be critical for the acceptor recognition by the enzyme. Entry 2, where the C2 hydroxyl group of Glc is replaced with an acetamido group, was not accepted by WciN, indicating WciN is specific to Glc with an equatorial C2 hydroxyl group. Moreover, the enzyme-lipid recognition may be an essential part of the acceptor recognition mechanism of WciN since UDP-Glc (entry 3) with a terminal uracil was unacceptable as an acceptor. Interestingly, the recognition of lipid moiety is quite relaxed since a saturated alkyl-based lipid with only 11 carbons is sufficient for the WciN activity.

Determination of divalent metal cation dependence of WciN

Two structural folds, named GT-A and GT-B, have been identified for GTs.[26] Most of GT-A GTs possess a DXD motif, by which they can coordinate a divalent metal cation, which in turn facilitates the departure of the nucleoside pyrophosphate leaving group during glycosyltransfer. In contrast, GT-B GTs were found to be divalent metal cation-independent, as they instead use the helix dipole effect or positively charged side chains to interact with the pyrophosphate of the donor.[26] In order to investigate the metal ion requirements of WciN, several divalent metal cations (Mg^{2+}, Mn^{2+}, or Zn^{2+}) were incubated in a series of reaction mixtures with or without the addition of EDTA. Results showed that WciN is divalent metal cation-dependent since the addition of a divalent metal cation is necessary for the activity of WciN and the removal of divalent metal cations (e.g. Mg^{2+}) by adding 10mM of EDTA into the reaction completely abolishes the WciN activity. The observed divalent metal cation-dependence suggests that WciN belongs to the GT-A superfamily.

Identification of the DXD motif of WciN

The divalent metal cation-dependence test of WciN indicated that WciN should contain a DXD motif or a variant of this motif. By analyzing the WciN sequence, three DXD motif candidates were found, including ^{28}DLD^{30}, ^{102}DSD^{104}, and ^{116}DID^{118}. To further identify which of them is involved in the binding of a divalent metal cation, six WciN mutants (D28A, D30A, D102A, D104A, D116A, and D118A) were prepared accordingly. The substitution of aspartic acid by alanine would eliminate the ability of WciN to bind a divalent metal cation and thus inactivate the enzyme. Six WciN mutants were applied to the established *in vitro* activity assay. Among the six mutants, two (D102A, D104A) lost activities while the other four remained active, which clearly indicates that the ^{102}DSD^{104} motif is critical for the WciN activity, and involved in the interaction of WciN with divalent metal cations.

A multiple sequence alignment was performed to further demonstrate the importance of the ^{102}DSD^{104} motif (Fig. 7). By searching the CAZY GT database (http://www.cazy.org), WciN is grouped into GT family 8. Ten bacterial GTs were randomly selected from this GT family, and together with WciN, were analyzed using multiple sequence alignment program (http://www.ebi.ac.uk/Tools/msa/clustalw2/). All of these GTs were found to contain a highly conserved DXD motif at the equivalent position of ^{102}DSD^{104}, which further confirms the importance of the ^{102}DSD^{104} motif to the WciN activity.

Discussion

In order to study the WciN function, we chemically synthesized Glcα-PP-O$(CH_2)_{10}CH_3$ for use as an acceptor substrate. The reasons to prompted us to test a synthetic acceptor with a

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

truncated lipid moiety are: (1) The natural acceptor of WciN is Glcα-PP-Und, which is produced by bacterial cells in a very small amount, making isolation impractical.[27] (2) Chemical synthesis may provide it on a relatively larger scale, but requires a very expensive undecaprenol precursor.[10,28] (3) Glcα-PP-Und has a 55-carbon hydrophobic lipid chain and is thus highly prone to aggregate, rendering it very difficult to handle owing to the poor solubility in enzymatic reaction solution.[29] Although Und-PP-linked substrates have been previously used in several studies for enzyme characterization, it should be noted that organic solvent (e.g. DMSO) and detergent (e.g. Triton X-100) must be supplemented to help dissolve the aggregates.[10,28,30] (4) Sequence analysis revealed that WciN does not contain any transmembrane segments, suggesting that the 55-carbon lipid may not be strictly required for the enzyme activity. In fact, it has been demonstrated that several bacterial GTs acting on Und-linked substrates are relaxed in their recognition of the lipid moiety.[31–34] We thus designed and chemically synthesized a Glcα-PP-lipid acceptor with an alkyl lipid chain of only 11 carbons. Using this acceptor mimic, WciN was characterized as an α-1,3-galactosyl transferase.

Two structurally similar substrates (GalNAcα-PP-$O(CH_2)_{11}$-O-Phenyl and GlcNAcα-PP-$O(CH_2)_{11}$-O-Phenyl) were previously synthesized to characterize the *E. coli* GTs.[32,33,35] In the present study, a similar synthetic scheme was initially employed to synthesize Glcα-PP-$O(CH_2)_{10}CH_3$. However, our initial experiments showed that, for lipid pyrophosphate derivatives of α-D-glucopyranosyl, the deacetylation of the peracetylated Glcα-PP-$O(CH_2)_{10}CH_3$ intermediate is difficult because of the extreme alkali-instability of the product. The reason could be that the C2 hydroxyl group of the glycosyl residue is equatorial and is sterically suited for the formation of 1,2-cyclic phosphate derivatives with cleavage of the pyrophosphate linkage in the alkaline reaction mixture (Fig. S1 of the Supporting Information). So upon the completion of the Zemplen deacetylation of peracetylated Glcα-PP-$O(CH_2)_{10}CH_3$, the cation-exchange resin (pyridinium form, week acidity) was immediately added to neutralize the reaction mixture. With this optimization, Glcα-PP-$O(CH_2)_{10}CH_3$ could be obtained with an overall yield >70%. Otherwise the final product would be degraded quickly upon concentration by rotovap.

WciN belongs to CAZY GT family 8, and members of this family employ retaining glycosyltranfer mechanism, in which glycosyl is transferred from the NDP-sugar donor to the acceptor with the retention of anomeric configuration. The structure analysis of the WciN disaccharide product by NMR revealed an α-1,3-linkage between Gal and Glc, consistent with the retaining glycosyltranfer mechanism. Furthermore, it was found that WciN, like those previously characterized bacterial GTs,[31–34] is also relaxed in the recognition of lipid moiety of the acceptor. The synthetic acceptor contains an 11-carbon alkyl lipid chain, quite different from the natural Und lipid in the saturation degree and the length. The retention of the WciN activity on such an acceptor suggests that the saturation degree and the length of lipid are not essential in determining the WciN activity. In future studies, simplified structures like this could thus be designed to replace complex natural substrates. Interestingly, aside from GTs, the relaxed specificity towards lipid moiety was also observed with other types of enzymes acting on Und-PP-linked glycan intermediates, such as O-antigen polymerase Wzy,[10] O-antigen ligase WaaL,[28] and peptidoglycan glycosyltransferases.[36] The reason behind this phenomenon remains unknown. To gain more insight into the mechanism of WciN for the acceptor lipid recognition, crystallization trials of the WciN-acceptor complex are undergoing in our laboratory.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Acknowledgments

Funding:

L.C. is grateful for the startup support from University of South Carolina and USC Salkehatchie. J.C. acknowledges the financial support from the National Natural Science Foundation of China Grant DE019451. P.G.W. acknowledges the financial support from National Institutes of Health Grant R01 GM085267.

We are grateful to Carlos J. Orihuela (the University of Texas Health Science Center at San Antonio) for providing the *S. pneumoniae* type 6B strain.

Abbreviations

COSY	correlation spectroscopy
DMSO	dimethyl sulfoxide
EDTA	ethylenediaminetetraacetic acid
Gal	galactose
GalNAc	*N*-Acetyl-D-galactosamine
Glc	glucose
GlcNAc	*N*-Acetyl-D-glucosamine
Glc-1-P	glucose-1-phosphate
HMBC	heteronuclear multiple bond correlation
HSQC	heteronuclear single quantum coherence
NDP	nucleoside diphosphate
NMR	nuclear magnetic resonance
PMSF	phenylmethylsulfonyl fluoride
Rha	rhamnose
Rib-ol	ribitol
UDP	uridine diphosphate

References

1. Neeleman C, Geelen SP, Aerts PC, Daha MR, Mollnes TE, Roord JJ, Posthuma G, van Dijk H, Fleer A. Resistance to both complement activation and phagocytosis in type 3 pneumococci is mediated by the binding of complement regulatory protein factor H. Infect Immun. 1999; 67:4517–4524. [PubMed: 10456894]
2. Saksouk N, Pelosi L, Colin-Morel P, Boumedienne M, Abdian PL, Geremia RA. The capsular polysaccharide biosynthesis of *Streptococcus pneumoniae* serotype 8: functional identification of the glycosyltransferase WciS (Cap8H). Biochem J. 2005; 389:63–72. [PubMed: 15766331]
3. Weinberger DM, Trzcinski K, Lu YJ, Bogaert D, Brandes A, Galagan J, Anderson PW, Malley R, Lipsitch M. Pneumococcal capsular polysaccharide structure predicts serotype prevalence. PLoS Pathog. 2009; 5:e1000476. [PubMed: 19521509]
4. Mavroidi A, Aanensen DM, Godoy D, Skovsted IC, Kaltoft MS, Reeves PR, Bentley SD, Spratt BG. Genetic relatedness of the *Streptococcus pneumoniae* capsular biosynthetic loci. J Bacteriol. 2007; 189:7841–7855. [PubMed: 17766424]
5. Bentley SD, Aanensen DM, Mavroidi A, Saunders D, Rabbinowitsch E, Collins M, Donohoe K, Harris D, Murphy L, Quail MA, Samuel G, Skovsted IC, Kaltoft MS, Barrell B, Reeves PR, Parkhill J, Spratt BG. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. PLoS Genet. 2006; 2:e31. [PubMed: 16532061]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

6. Valvano MA. Export of O-specific lipopolysaccharide. Front Biosci. 2003; 8:s452–471. [PubMed: 12700099]
7. Patel KB, Furlong SE, Valvano MA. Functional analysis of the C-terminal domain of the WbaP protein that mediates initiation of O antigen synthesis in *Salmonella enterica*. Glycobiology. 2010; 20:1389–1401. [PubMed: 20591829]
8. Xayarath B, Yother J. Mutations blocking side chain assembly, polymerization, or transport of a Wzy-dependent *Streptococcus pneumoniae* capsule are lethal in the absence of suppressor mutations and can affect polymer transfer to the cell wall. J Bacteriol. 2007; 189:3369–3381. [PubMed: 17322316]
9. Cartee RT, Forsee WT, Bender MH, Ambrose KD, Yother J. CpsE from type 2 *Streptococcus pneumoniae* catalyzes the reversible addition of glucose-1-phosphate to a polyprenyl phosphate acceptor, initiating type 2 capsule repeat unit formation. J Bacteriol. 2005; 187:7425–7433. [PubMed: 16237026]
10. Woodward R, Yi W, Li L, Zhao G, Eguchi H, Sridhar PR, Guo H, Song JK, Motari E, Cai L, Kelleher P, Liu X, Han W, Zhang W, Ding Y, Li M, Wang PG. *In vitro* bacterial polysaccharide biosynthesis: defining the functions of Wzy and Wzz. Nat Chem Biol. 2010; 6:418–423. [PubMed: 20418877]
11. Islam ST, Gold AC, Taylor VL, Anderson EM, Ford RC, Lam JS. Dual conserved periplasmic loops possess essential charge characteristics that support a catch-and-release mechanism of O-antigen polymerization by Wzy in *Pseudomonas aeruginosa* PAO1. J Biol Chem. 2011; 286:20600–20605. [PubMed: 21498511]
12. Guo H, Yi W, Song JK, Wang PG. Current understanding on biosynthesis of microbial polysaccharides. Curr Top Med Chem. 2008; 8:141–151. [PubMed: 18289083]
13. Liu D, Cole RA, Reeves PR. An O-antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase. J Bacteriol. 1996; 178:2102–2107. [PubMed: 8606190]
14. Feldman MF, Marolda CL, Monteiro MA, Perry MB, Parodi AJ, Valvano MA. The activity of a putative polyisoprenol-linked sugar translocase (Wzx) involved in *Escherichia coli* O antigen assembly is independent of the chemical structure of the O repeat. J Biol Chem. 1999; 274:35129–35138. [PubMed: 10574995]
15. Bengoechea JA, Pinta E, Salminen T, Oertelt C, Holst O, Radziejewska-Lebrecht J, Piotrowska-Seget Z, Venho R, Skurnik M. Functional characterization of Gne (UDP-N-acetylglucosamine-4-epimerase), Wzz (chain length determinant), and Wzy (O-antigen polymerase) of *Yersinia enterocolitica* serotype O:8. J Bacteriol. 2002; 184:4277–4287. [PubMed: 12107146]
16. Grozdanov L, Zahringer U, Blum-Oehler G, Brade L, Henne A, Knirel YA, Schombel U, Schulze J, Sonnenborn U, Gottschalk G, Hacker J, Rietschel ET, Dobrindt U. A single nucleotide exchange in the *wzy* gene is responsible for the semirough O6 lipopolysaccharide phenotype and serum sensitivity of *Escherichia coli* strain Nissle 1917. J Bacteriol. 2002; 184:5912–5925. [PubMed: 12374825]
17. Poolman JT. Pneumococcal vaccine development. Expert Rev Vaccines. 2004; 3:597–604. [PubMed: 15485339]
18. Alexander J, del Guercio MF, Frame B, Maewal A, Sette A, Nahm MH, Newman MJ. Development of experimental carbohydrate-conjugate vaccines composed of *Streptococcus pneumoniae* capsular polysaccharides and the universal helper T-lymphocyte epitope (PADRE). Vaccine. 2004; 22:2362–2367. [PubMed: 15193395]
19. Lee CJ, Lee LH, Frasch CE. Protective immunity of pneumococcal glycoconjugates. Crit Rev Microbiol. 2003; 29:333–349. [PubMed: 14636043]
20. Palcic MM. Glycosyltransferases as biocatalysts. Curr Opin Chem Biol. 2011; 15:226–233. [PubMed: 21334964]
21. Aanensen DM, Mavroidi A, Bentley SD, Reeves PR, Spratt BG. Predicted functions and linkage specificities of the products of the *Streptococcus pneumoniae* capsular biosynthetic loci. J Bacteriol. 2007; 189:7856–7876. [PubMed: 17766420]
22. Kolkman MA, van der Zeijst BA, Nuijten PJ. Functional analysis of glycosyltransferases encoded by the capsular polysaccharide biosynthesis locus of *Streptococcus pneumoniae* serotype 14. J Biol Chem. 1997; 272:19502–19508. [PubMed: 9235953]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

23. Kenne L, Lindberg B, Madden JK. Structural studies of the capsular antigen from *Streptococcus pneumoniae* type 26. Carbohydr Res. 1979; 73:175–182. [PubMed: 38908]
24. Rebers PA, Heidelberger M. The specific polysaccharide of type VI pneumococcus. II The repeating unit. J Am Chem Soc. 1961; 83:3056–3059.
25. Guo H, Lokko K, Zhang Y, Yi W, Wu Z, Wang PG. Overexpression and characterization of Wzz of *Escherichia coli* O86:H2. Protein Expr Purif. 2006; 48:49–55. [PubMed: 16603378]
26. Lairson LL, Henrissat B, Davies GJ, Withers SG. Glycosyltransferases: structures, functions, and mechanisms. Annu Rev Biochem. 2008; 77:521–555. [PubMed: 18518825]
27. Whitfield C. Glycan chain-length control. Nat Chem Biol. 2010; 6:403–404. [PubMed: 20479750]
28. Han W, Wu B, Li L, Zhao G, Woodward R, Pettit N, Cai L, Thon V, Wang PG. Defining function of lipopolysaccharide O-antigen ligase WaaL using chemoenzymatically synthesized substrates. J Biol Chem. 2012; 287:5357–5365. [PubMed: 22158874]
29. Ye XY, Lo MC, Brunner L, Walker D, Kahne D, Walker S. Better substrates for bacterial transglycosylases. J Am Chem Soc. 2001; 123:3155–3156. [PubMed: 11457035]
30. Glover KJ, Weerapana E, Numao S, Imperiali B. Chemoenzymatic synthesis of glycopeptides with PglB, a bacterial oligosaccharyl transferase from Campylobacter jejuni. Chem Biol. 2005; 12:1311–1315. [PubMed: 16356848]
31. Riley JG, Menggad M, Montoya-Peleaz PJ, Szarek WA, Marolda CL, Valvano MA, Schutzbach JS, Brockhausen I. The *wbbD* gene of *E. coli* strain VW187 (O7:K1) encodes a UDP-Gal: GlcNAc{alpha}-pyrophosphate-R {beta}1,3-galactosyltransferase involved in the biosynthesis of O7-specific lipopolysaccharide. Glycobiology. 2005; 15:605–613. [PubMed: 15625181]
32. Brockhausen I, Hu B, Liu B, Lau K, Szarek WA, Wang L, Feng L. Characterization of two beta-1,3-glucosyltransferases from *Escherichia coli* serotypes O56 and O152. J Bacteriol. 2008; 190:4922–4932. [PubMed: 18487334]
33. Xu C, Liu B, Hu B, Han Y, Feng L, Allingham JS, Szarek WA, Wang L, Brockhausen I. Biochemical characterization of UDP-Gal:GlcNAc-pyrophosphate-lipid beta-1,4-Galactosyltransferase WfeD, a new enzyme from *Shigella boydii* type 14 that catalyzes the second step in O-antigen repeating-unit synthesis. J Bacteriol. 2011; 193:449–459. [PubMed: 21057010]
34. Brown S, Zhang YH, Walker S. A revised pathway proposed for *Staphylococcus aureus* wall teichoic acid biosynthesis based on in vitro reconstitution of the intracellular steps. Chem Biol. 2008; 15:12–21. [PubMed: 18215769]
35. Yi W, Yao Q, Zhang Y, Motari E, Lin S, Wang PG. The *wbnH* gene of *Escherichia coli* O86:H2 encodes an alpha-1,3-N-acetylgalactosaminyl transferase involved in the O-repeating unit biosynthesis. Biochem Biophys Res Commun. 2006; 344:631–639. [PubMed: 16630548]
36. Perlstein DL, Wang TS, Doud EH, Kahne D, Walker S. The role of the substrate lipid in processive glycan polymerization by the peptidoglycan glycosyltransferases. J Am Chem Soc. 2010; 132:48–49. [PubMed: 20017480]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

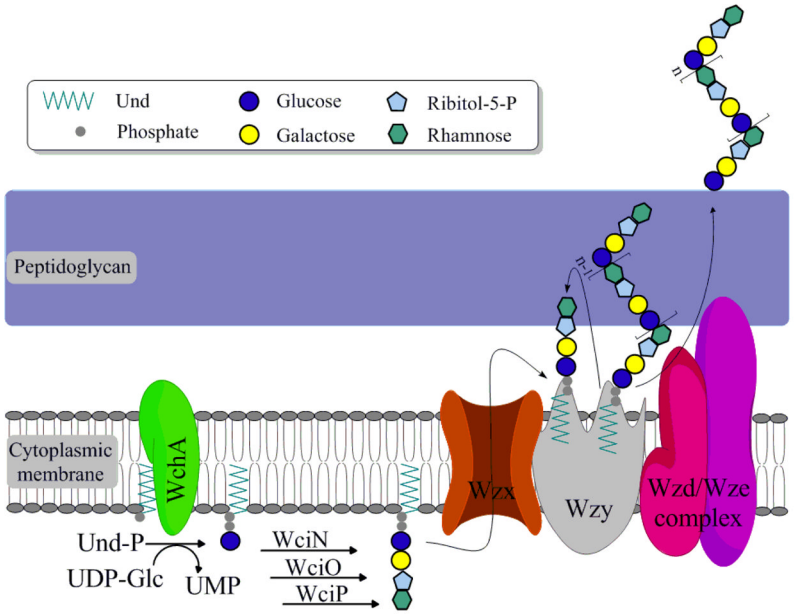


Fig. 1.

The current model for the Wzy-dependent pneumococcal CPS biosynthesis and the export to the cell surface. The type 6B CPS pathway is presented here as an example.

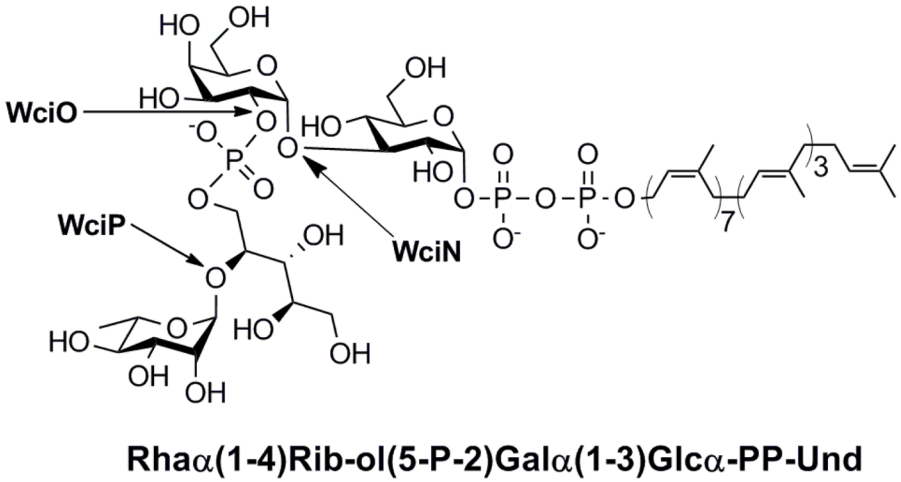


Fig. 2.
The structure of the pneumococcus type 6B CPS RU and the proposed activities of three GTs.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

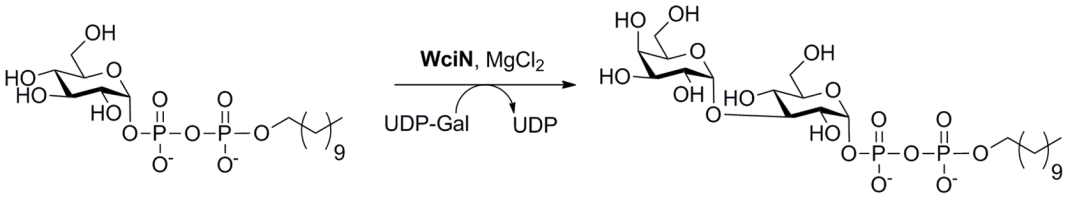


Fig. 3.

In vitro activity characterization of the WciN activity using a chemically synthesized acceptor: Glcα-PP-O$(CH_2)_{10}CH_3$.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

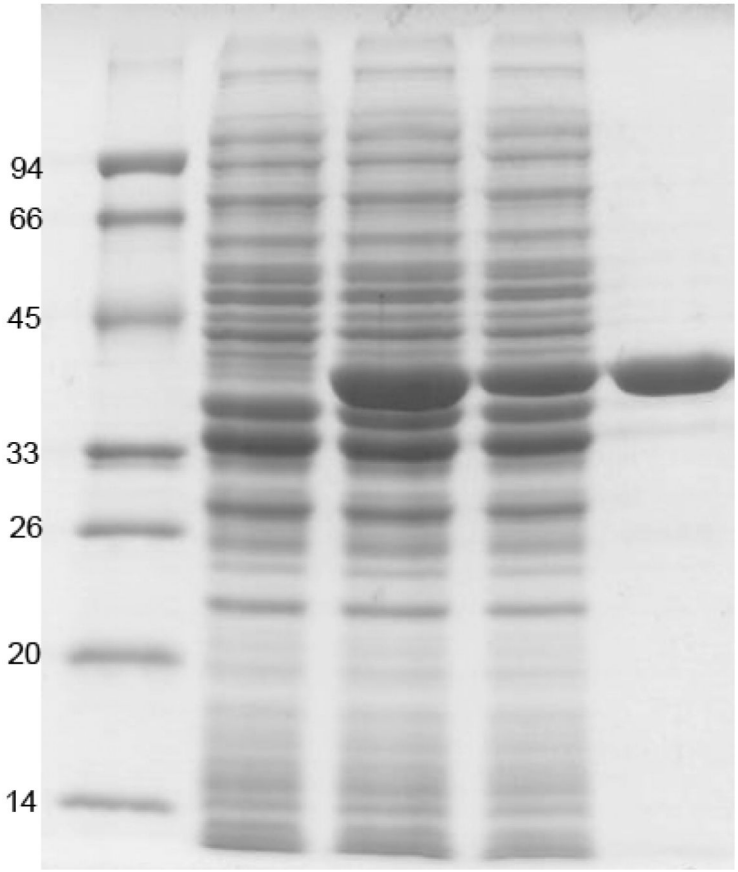


Fig. 4.

The expression and purification of WciN. Lane 1, Protein molecular weight standard; Lane 2, whole *E. coli* BL21(DE3) cells with empty plasmid pET-15b; Lane 3, whole cells with pET-15b-*wicN* induced by IPTG; Lane 4, supernatant after the cell disruption; Lane 5, purified WciN using Ni-affinity column.

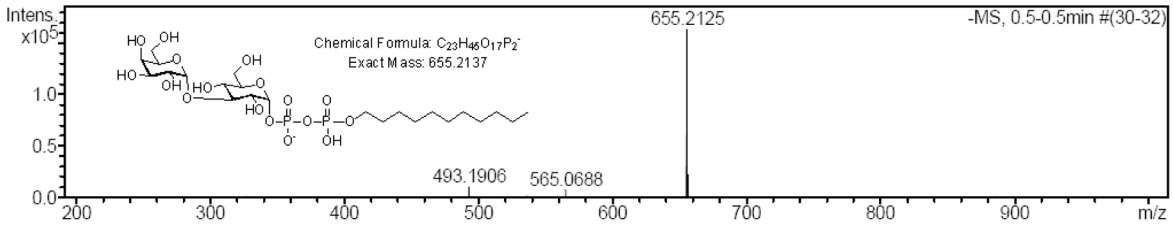


Fig. 5.
HRMS analysis of Gala1,3Glcα-PP-O$(CH2)_{10}CH_3$.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

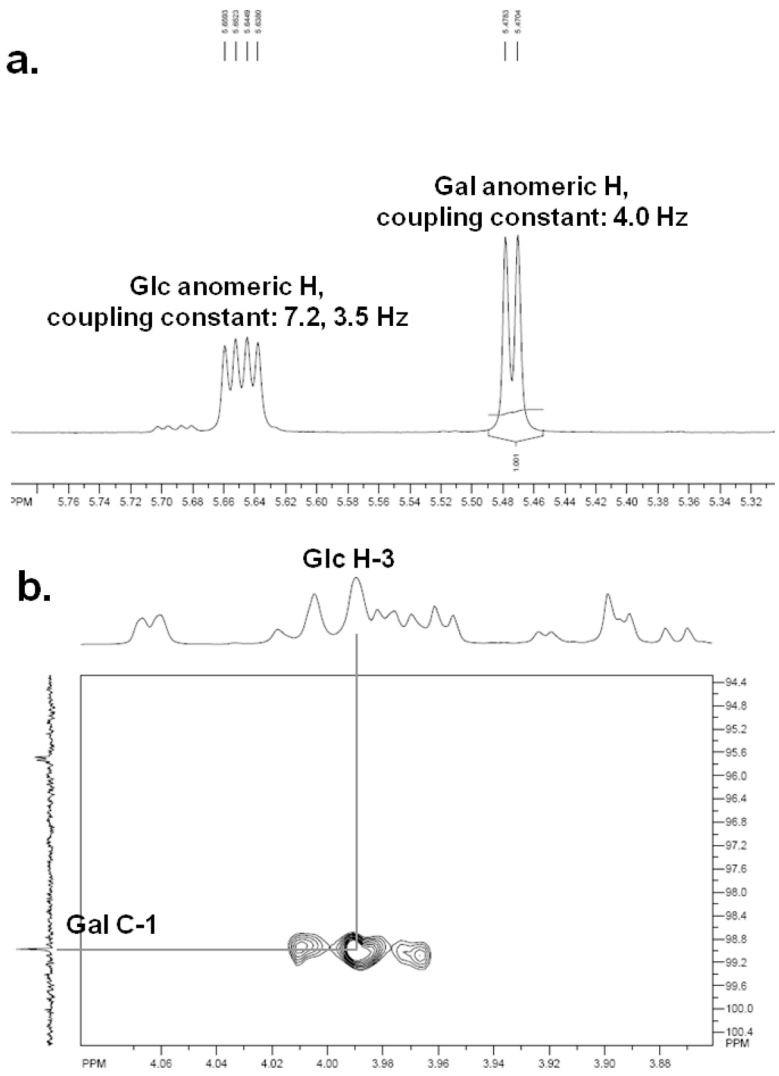


Fig. 6.

^{1}H NMR and HMBC spectra of Galα1,3Glcα-PP-O(CH2)$_{10}$CH$_3$.

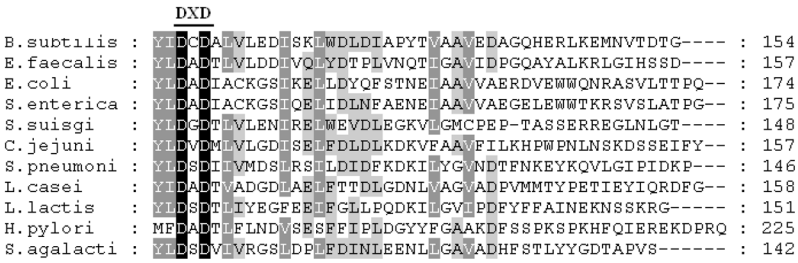


Fig. 7.

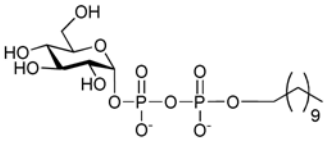
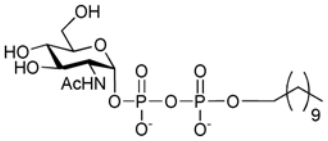
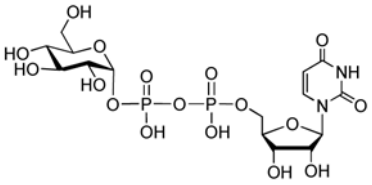
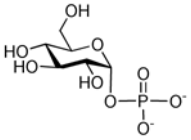
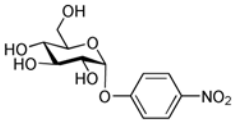
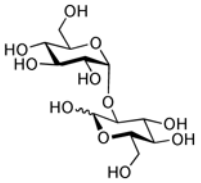
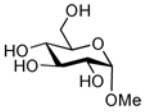
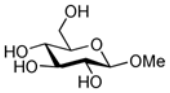
Multiple sequence alignment analysis of GTs from GT family 8. Genbank accession numbers for the GTs are as follows: *B. subtilis*, BAI87514; *E. faecalis*, ADX80404; *E. coli*, CBG36759; *S. enterica*, CBW19778; *S. suisgi*, ADV70317; *C. jejuni*, CAI38726; *S. pneumoniae*, AAL68409; *L. casei*, CAQ66284; *L. lactis*, CAL98281; *H. pylori*, ADI34271; *S. agalactiae*, AE014258.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

Investigation on the acceptor substrate specificity of WciN

No.	Acceptor	Structure	Activity*
(1)	Glcα-PP-O$(CH2)_{10}CH_3$		+
(2)	GlcNAcα-PP-O$(CH2)_{10}CH_3$		−
(3)	UDP-Glc		−
(4)	Glcα-P		−
(5)	4-Nitrophenyl α-D-glucopyranoside		−
(6)	Glcα(1-2)Glc		−
(7)	Glcα-OMe		−
(8)	Glcβ-OMe		−

*+: The WciN activity was observed; −: No detectable activity.