Supplemental data

GENERAL MATERIALS AND PROCEDURE FOR SYNTHESIS OF GalNAc/GlcNAc-PP-LIPID

All solvents were dried with a solvent-purification system from Innovative Technology, Inc. All reagents were obtained from commercial sources and used without further purification. The 200-400 mesh silica gel 60 RP-18 was utilized for purification. ^{1}H, ^{13}C and ^{31}P NMR spectra were recorded at the indicated field strengths. Mass spectral data was collected using a Shimadzu LCMS-2010A Liquid Chromatography Mass Spectrometer or a Bruker micrOTOF Instrument.

The general procedure for chemical synthesis of GalNAc/GlcNAc-PP-lipid compounds is shown as Fig. S1. Specifically, the N,N'-Carbonyldiimidazole (CDI) (23 mg, 0.14 mmol) was added to a reaction vessel containing the diisopropylethylammonium salt of GalNAc/GlcNAc-1-PO_4^{2-} (23 mg, 0.034 mmol). The vessel was flushed with argon, after which anhydrous THF was added (3 mL). The reaction was stirred at room temperature for 2 h. Addition of dry CH_3OH (42 μL) followed by stirring for an additional 1 h at room temperature served to remove excess CDI. The reaction mixture was concentrated and then anhydrous THF (3 mL) was once again added. To a separate reaction vessel, lipid phosphate (0.026 mmol) was added, and the vessel was then flushed with argon. The activated GalNAc/GlcNAc-1-PO_4^{2-} was transferred into this vessel via syringe. The reaction was stirred at room temperature for 3 days, after which the reaction mixture was concentrated. Dissolution of the residue in water, followed by filtration through a short bed of C-18 reverse phase silica gel (water then 1:1 Water/Isopropanol) afforded the crude product.

The crude product was treated with a 0.1% $NaOCH_3$ in CH_3OH solution (3 mL). The reaction was stirred at room temperature for 40 min. The reaction mixture was neutralized with Amberlyst-15 ion exchange resin until pH around 7 and then concentrated, after which the residue was dissolved in water. Purification via C-18 reverse phase column chromatography (Solvent A: Isopropanol (Lipids: Undecaprenyl, Solanesyl,) or CH_3CN (Lipids: Heptaprenyl, Pentaprenyl, MS-Pentaprenyl, cis-Farnesyl, Undecyl, Geranyl); Solvent B = 1.7% NH_4HCO_3; A/B = 0%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45% and 50% (10 mL each)) provided GalNAc/GlcNAc-PP-lipid as a white solid following lyophilization.

SYNTHESIS AND CHARACTERIZATION OF EACH OF GalNAc/GlcNAc-PP-LIPID COMPOUNDS

Chemical synthesis and characterization of some GalNAc/GlcNAc-PP-lipid compounds have been reported in our previously published paper (Woodward R, *et al*. (2010) *In vitro*

bacterial polysaccharide biosynthesis: defining the functions of Wzy and Wzz. *Nat Chem Biol* 6:418-423.) and those data as well as the data for the newly-synthesized compounds in this study were all listed below for reference.

Chemical synthesis of GalNAc-PP-Geranyl-GalNAc-PP-Geranyl was prepared using the general procedure as described above. Following the deprotection step, the crude product was purified via C-18 reverse phase column chromatography to afford GalNAc-PP-Geranyl as a white solid (10 mg, 11% over two steps). ^{1}H NMR (500 MHz, D_2O): δ 5.57 (dd, *J* = 3.4 Hz, *J* = 7.0 Hz, 1H), 5.48 (td, *J* = 1.1 Hz, *J* = 7.1 Hz, 1H), 5.28-5.22 (m, 1H), 4.53 (t, *J* = 6.9 Hz, 2H), 4.29 (dt, *J* = 3.0 Hz, *J* = 10.9 Hz, 1H), 4.23 (t, *J* = 6.2 Hz, 1H), 4.08 (d, *J* = 2.6 Hz, 1H), 4.01 (dd, *J* = 3.2 Hz, *J* = 10.9 Hz, 1H), 3.85-3.78 (m, 2H), 2.23-2.17 (m, 2H), 2.16-2.12 (m, 2H), 2.11 (s, 3H), 1.76 (s, 3H), 1.73 (s, 3H), 1.67 (s, 3H); ^{13}C NMR (125 MHz, D_2O): δ 160.3, 143.0, 133.7, 124.1, 121.6, 94.7, 72.1, 68.4, 67.8, 62.5, 61.1, 49.8, 38.8, 25.6, 24.8, 22.2, 17.0, 15.6; ^{31}P NMR (162 MHz, D_2O): δ -10.2, -12.5; LRMS (*m/z*): $[M-H]^-$ calcd. for $C_{18}H_{32}NO_{12}P_2$, 516.1; found 516.1.

Chemical synthesis of GalNAc-PP-cis-Farnesyl-GalNAc-PP-cis-Farnesyl was prepared using the general procedure as described above. Following the deprotection step, the crude product was purified via C-18 reverse phase column chromatography to afford GalNAc-PP-cis-Farnesyl as a white solid (11 mg, 6% over two steps). ^{1}H NMR (500 MHz, D_2O): δ 5.58 (dd, *J* = 3.4 Hz, *J* = 7.0 Hz, 1H), 5.51 (t, *J* = 6.1 Hz, 1H), 5.32-5.22 (m, 2H), 4.52 (t, *J* = 7.0 Hz, 2H), 4.31 (dt, *J* = 3.0 Hz, *J* = 10.9 Hz, 1H), 4.25 (t, *J* = 6.2 Hz, 1H), 4.09 (d, *J* = 3.0 Hz, 1H), 4.01 (dd, *J* = 3.2 Hz, *J* = 10.9 Hz, 1H), 3.88-3.75 (m, 2H), 2.24-2.13 (m, 8H), 2.12 (s, 3H), 1.83 (s, 3H), 1.75 (s, 6H), 1.69 (s, 3H); ^{13}C NMR (125 MHz, D_2O): δ 175.0, 142.8, 137.2, 133.7, 124.8, 124.4, 120.5, 94.7, 72.1, 68.5, 67.9, 62.9, 61.1, 49.8, 32.1, 31.6, 31.2, 25.9, 24.9, 22.7, 22.5, 22.2, 17.0; ^{31}P NMR (162 MHz, D_2O): δ -10.2, -12.5; LRMS (*m/z*): $[M-H]^-$ calcd. for $C_{23}H_{40}NO_{12}P_2$, 584.2; found 584.2.

Chemical synthesis of GalNAc-PP-MS-Pentaprenyl-GalNAc-PP-MS-Pentaprenyl was prepared using the general procedure as described above. Following the deprotection step, the crude product was purified via C-18 reverse phase column chromatography to afford GalNAc-PP-MS-Pentaprenyl as a white solid (27 mg from 65 mg lipid phosphate, 26% over 2 steps). ^{1}H NMR (500 MHz, D_2O): δ 5.55 (dd, *J* = 3.3 Hz, *J* = 7.1 Hz, 1H), 5.22-5.05 (m, 4H), 4.26 (dt, *J* = 2.9 Hz, *J* = 11.0 Hz, 1H), 4.19 (dd, *J* = 5.2 Hz, *J* = 7.1 Hz, 1H), 4.06-3.93 (m, 4H), 3.84-3.72 (m, 2H), 2.11-2.02 (m, 11H), 2.01-1.92 (m, 6H), 1.69 (s, 3H), 1.66 (s, 3H) 1.61 (s, 3H), 1.58 (s, 6H), 1.54-1.24 (m, 4H), 1.22-1.10 (m, 1H), 0.92 (d, *J* = 6.5 Hz, 3H); ^{13}C NMR (125 MHz, D_2O): δ 174.5, 135.0, 134.8, 134.5, 130.8, 125.7, 124.4, 124.2, 94.7, 72.1, 68.6, 67.7, 64.9, 61.3, 49.9, 49.8, 39.6, 37.3, 37.1, 31.8, 29.0, 26.6, 26.6, 26.5, 25.4, 24.9, 23.2, 22.3, 18.8, 18.8, 17.4, 15.7, 15.7; ^{31}P

NMR (162 MHz, D_2O): δ -9.9, -11.8; LRMS (*m/z*): $[M-H]^-$ calcd. for $C_{33}H_{58}NO_{12}P_2$, 722.3; found, 722.3.

Chemical synthesis of GalNAc-PP-Pentaprenyl-GalNAc-PP-Pentaprenyl was prepared using the general procedure as described above. Following the deprotection step, the crude product was purified via C-18 reverse phase column chromatography to afford GalNAc-PP-Pentaprenyl as a white solid (22 mg from 28 mg lipid phosphate, 49% over 2 steps). ^1H NMR (500 MHz, CD_3OD): δ 5.67 (dd, *J* = 3.1 Hz, *J* = 6.8 Hz, 1H), 5.56 (t, *J* = 6.6 Hz, 1H), 5.30-5.16 (m, 4H), 4.58 (t, *J* = 6.5 Hz, 2H), 4.36 (d, *J* = 10.9 Hz, 1H), 4.30 (dd, *J* = 6.0 Hz, 1H), 4.14 (d, *J* = 3.0 Hz, 1H), 4.07 (dd, *J* = 2.9 Hz, *J* = 10.9 Hz, 1H), 3.94-3.84 (m, 2H), 2.30-2.23 (m, 2H), 2.22-2.11 (m, 13H), 2.10-2.02 (m, 4H), 1.85 (s, 3H), 1.80 (s, 3H), 1.75 (s, 3H), 1.71 (s, 3H), 1.68 (s, 6H); ^{13}C NMR (125 MHz, D_2O): δ 174.7, 160.3, 141.2, 135.8, 134.9, 134.4, 130.6, 124.6, 124.5, 124.3, 121.2, 94.7, 72.2, 68.7, 67.7, 62.7, 61.3, 49.9, 48.9, 39.7, 39.7, 32.0, 31.9, 26.7, 26.6, 26.1, 25.4, 23.1, 22.9, 22.3, 17.4, 15.8, 15.7; ^{31}P NMR (162 MHz, D_2O): δ -10.4, -13.1; LRMS (*m/z*): $[M-H]^-$ calcd. for $C_{33}H_{56}NO_{12}P_2$, 720.3; found, 720.3.

Chemical synthesis of GalNAc-PP-Heptaprenyl-GalNAc-PP-Heptaprenyl was prepared using the general procedure as described above. Following the deprotection step, the crude product was purified via C-18 reverse phase column chromatography to afford GalNAc-PP-Heptaprenyl as a white solid (22 mg from 42 mg lipid phosphate, 36% over 2 steps). ^1H NMR (500 MHz, D_2O): δ 5.61 (dd, *J* = 2.7 Hz, *J* = 5.9 Hz, 1H), 5.50 (t, *J* = 5.5 Hz, 1H), 5.24-5.08 (m, 6H), 4.51 (t, *J* = 5.7 Hz, 2H), 4.29 (d, *J* = 10.8 Hz, 1H), 4.23 (t, *J* = 5.5 Hz, 1H), 4.07 (d, *J* = 1.4 Hz, 1H), 4.00 (dd, *J* = 2.2 Hz, *J* = 10.7 Hz, 1H), 3.88-3.75 (m, 2H), 2.20-1.96 (m, 27H), 1.79 (s, 3H), 1.72 (s, 3H), 1.70 (s, 3H), 1.68 (s, 3H), 1.62 (s, 3H), 1.60 (s, 9H); ^{13}C NMR (125 MHz, D_2O): δ 174.5, 172.4, 172.2, 171.9, 140.8, 136.3, 135.9, 135.9, 135.9, 135.8, 132.1, 125.7, 125.7, 125.6, 125.4, 122.7, 122.6, 96.5, 70.6, 68.7, 64.0, 64.0, 62.3, 41.0, 41.0, 41.0, 40.9, 40.8, 34.5, 33.2, 30.9, 30.8, 30.7, 30.6, 30.5, 28.0, 27.9, 27.8, 27.7, 27.7, 27.7, 26.0, 24.0, 23.8, 23.0, 20.8, 20.7, 20.7, 17.9, 16.8, 16.2, 15.7, 14.5; 174.7, 165.3, 134.8, 134.6, 134.3, 130.4, 125.2, 125.0, 124.8, 124.5, 124.3, 124.3, 121.2, 94.6, 72.1, 68.7, 67.7, 62.6, 61.3, 49.9, 39.7, 39.7, 32.0, 31.9, 31.9, 29.9, 29.8, 29.7, 29.5, 26.7, 26.6, 26.2, 25.4, 23.3, 23.2, 23.0, 23.0, 22.3, 17.3, 15.7; ^{31}P NMR (162 MHz, D_2O): δ -9.7, -11.4; LRMS (*m/z*): $[M-H]^-$ calcd. for $C_{43}H_{72}NO_{12}P_2$, 856.5; found, 856.7.

Chemical synthesis of GalNAc-PP-Solanesyl-GalNAc-PP-Solanesyl was prepared using the general procedure as described above. However, the two purification steps were reversed as the final product proved too insoluble to obtain suitable spectral data. The crude product was thus purified via C-18 reverse phase column chromatography to afford Peracetylated GalNAc-PP-

Solanesyl as a white solid (35 mg from 45 mg lipid phosphate, 50%). ^{1}H NMR (500 MHz, CD_3OD): δ 5.61 (dd, *J* = 3.3 Hz, *J* = 7.3 Hz, 1H), 5.43 (t, *J* = 6.6 Hz, 2H), 5.22 (dd, *J* = 3.1 Hz, *J* = 11.3 Hz, 1H), 5.14-5.04 (m, 8H), 4.59 (t, *J* = 7.0 Hz, 1H), 4.53 (t, *J* = 6.3 Hz, 3H), 4.21 (dd, *J* = 8.2 Hz, *J* = 10.8 Hz, 1H), 4.04 (dd, *J* = 5.8 Hz, *J* = 10.8 Hz, 1H), 2.11 (s, 3H), 2.10-2.02 (m, 18H), 1.99 (s, 3H), 1.99-1.94 (m, 17H), 1.90 (s, 3H), 1.69 (s, 3H), 1.65 (s, 3H), 1.59 (s, 3H), 1.58 (s, 21H); ^{13}C NMR (125 MHz, CD_3OD): δ 174.5, 172.4, 172.2, 171.9, 140.8, 136.3, 135.9, 135.9, 135.9, 135.8, 132.1, 125.7, 125.7, 125.6, 125.4, 122.7, 122.6, 96.5, 70.6, 68.7, 64.0, 64.0, 62.3, 41.0, 41.0, 41.0, 40.9, 40.8, 34.5, 33.2, 30.9, 30.8, 30.7, 30.6, 30.5, 28.0, 27.9, 27.8, 27.7, 27.7, 27.7, 26.0, 24.0, 23.8, 23.0, 20.8, 20.7, 20.7, 17.9, 16.8, 16.2, 15.7, 14.5; ^{31}P NMR (162 MHz, CD_3OD): δ -9.6, -12.5; LRMS (*m/z*): $[M-H]^-$ calcd. for $C_{59}H_{94}NO_{15}P_2$, 1118.6; found, 1118.5.

Final deprotection with 0.1% $NaOCH_3$ in CH_3OH and filtration through a short bed of C-18 reverse phase silica gel afforded GalNAc-PP-Solanesyl (31 mg, 99%). LRMS (*m/z*): $[M-H]^-$ calcd. for $C_{53}H_{88}NO_{12}P_2$, 992.6; found 992.6.

Chemical synthesis of GalNAc-PP-Undecaprenyl-GalNAc-PP-Undecaprenyl was prepared using the general procedure as described above. However, the two purification steps were reversed as the final product proved too insoluble to obtain suitable spectral data. The crude product was thus purified via C-18 reverse phase column chromatography to afford Peracetylated GalNAc-PP-Undecaprenyl as a white solid (20 mg from 23 mg lipid phosphate, 59%). ^{1}H NMR (500 MHz, CD_3OD): δ 5.60 (dd, *J* = 3.2 Hz, *J* = 7.3 Hz, 1H), 5.42 (t, *J* = 6.1 Hz, 2H), 5.21 (dd, *J* = 3.2 Hz, *J* = 11.3 Hz, 1H), 5.14-5.05 (m, 10H), 4.58 (t, *J* = 6.9 Hz, 1H), 4.54-4.45 (m, 3H), 4.21 (dd, *J* = 8.1 Hz, *J* = 10.9 Hz, 1H), 4.04 (dd, *J* = 5.8 Hz, *J* = 10.8 Hz, 1H), 2.10 (s, 3H), 2.08-2.00 (m, 34H), 1.98 (s, 3H), 1.97 (s, 3H), 1.96-1.92 (m, 6H), 1.89 (s, 3H), 1.71 (s, 3H), 1.66-1.62 (m, 21H), 1.58 (s, 3H), 1.57 (s, 9H); ^{13}C NMR (125 MHz, CD_3OD): δ 174.4, 172.3, 172.2, 171.9, 140.6, 136.5, 136.4, 136.3, 136.3, 136.1, 135.9, 132.1, 126.3, 126.3, 126.3, 126.0, 125.6, 125.6, 125.6, 123.6, 123.6, 96.5, 70.4, 68.8, 68.7, 63.8, 63.8, 62.3, 41.0, 41.0, 40.9, 34.4, 33.4, 33.4, 33.4, 33.2, 33.0, 30.8, 30.8, 30.7, 30.6, 30.5, 30.5, 30.4, 28.0, 27.8, 27.8, 27.8, 27.7, 27.7, 27.7, 26.1, 24.0, 23.9, 23.8, 23.0, 20.8, 20.8, 20.7, 17.9, 16.3, 14.5, 14.0; ^{31}P NMR (162 MHz, CD_3OD): δ -9.7, -12.6; LRMS (*m/z*): $[M-H]^-$ calcd. for $C_{69}H_{110}NO_{15}P_2$, 1254.7; found, 1254.7.

Final deprotection with 0.1% $NaOCH_3$ in CH_3OH and filtration through a short bed of C-18 reverse phase silica gel afforded GalNAc-PP-Undecaprenyl (18 mg, 99%). LRMS (*m/z*): $[M-H]^-$ calcd. for $C_{63}H_{104}NO_{12}P_2$,1128.7; found, 1128.7.

Chemical synthesis of GlcNAc-PP-Pentaprenyl-GlcNAc-PP-Pentaprenyl was prepared in a pure form, and directly used without any purification after deacetylation of AcGlcNAc-PP-Pentaprenyl, which was purified by C-18 reverse phase column chromatography [Solvent A:

CH_3CN; Solvent B = 1.7% NH_4HCO_3; A/B = 0%→60% with 5% increment each time in concentration of A (10 mL each)]. Fraction containing product was lyophilized to afford pure AcGlcNAc-PP- Pentaprenyl as a white solid (14 mg from 40 mg lipid phosphate 21% yield). ^{1}H NMR (500 MHz, MeOD): δ 5.60 (br s, 1H), 5.48 (t, *J* = 6.5 Hz, 1H), 5.34 (t, *J* = 10.0 Hz, 1H), 5.17-5.09 (m, 5H), 4.56 (t, *J* = 6.5 Hz, 2H), 4.40-4.31 (m, 3H), 4.18 (dd, *J* = 1.5, 12.5 Hz, 1H), 2.16-1.97 (m, 8H), 2.07 (s, 3H), 2.01 (s, 3H), 2.00 (s, 3H), 1.97 (s, 3H), 1.76 (s, 3H), 1.70 (s, 3H), 1.68 (s, 3H), 1.63 (s, 3H), 1.61 (s, 3H); ^{13}C NMR (125 MHz, MeOD): δ 172.6, 171.1, 170.5, 169.9, 139.7, 135.2, 134.7, 134.5, 130.7, 124.4, 124.1, 121.7, 121.6, 94.5, 71.5, 68.6, 68.3, 62.6, 61.4, 39.4, 31.9, 31.6, 26.4, 26.21, 26.19, 24.5, 22.3, 21.4, 19.30, 19.26, 19.25, 16.4, 14.8; ^{31}P NMR (162 MHz, MeOD): δ -10.5, -13.4; LRMS (*m/z*): $[M-H]^-$ calcd. for $C_{39}H_{62}NO_{15}P_2$, 846.3, found 846.3.

Final deprotection with catalytic amount of 0.1% $NaOCH_3$ in CH_3OH and subsequent addition of Amberlyst-15 ion exchange resin to the reaction mixture until around neutral pH (pH 7.0) afforded the GlcNAc-PP-Pentaprenyl in quantitative yield. LRMS (*m/z*): $[M-H]^-$ calcd. for $C_{33}H_{56}NO_{12}P_2$, 720.2, found 720.2.

Chemical synthesis of GalNAc-PP-Undecyl-GalNAc-PP-Undecyl was prepared in a pure form, and directly used without any purification after deacetylation of AcGalNAc-PP-Undecyl, which was purified by C-18 reverse phase column chromatography [Solvent A: CH_3CN; Solvent B = 1.7% NH_4HCO_3; A/B = 0%→10% with 1% increment in concentration of A (10 mL each)]. Fraction containing product was lyophilized to afford pure AcGalNAc-PP-Undecyl as a white solid (19 mg from 30 mg lipid phosphate 30% yield). ^{1}H NMR (500 MHz, MeOD): δ 7.58 (s, 1H), 5.67-5.65 (m, 1H), 5.46 (d, *J* = 3.0 Hz, 1H), 5.24 (dd, *J* = 3.5, 11.0 Hz, 1H), 4.61-4.55 (m, 2H), 4.23 (dd, J = 3.0, 11.0 Hz, 1H), 4.09-4.06 (m, 1H), 4.03-3.99 (m, 2H), 2.15 (s, 3H), 2.023 (s, 3H), 2.021 (s, 3H), 1.95 (s, 3H), 1.70-1.64 (m, 2H), 1.41-1.31 (m, 16H), 0.92 (t, *J* = 7.0 Hz, 3H); ^{13}C NMR (125 MHz, MeOD): δ 172.8, 170.8, 170.7, 170.5, 119.0, 95.1(d, *J* = 5.0 Hz), 68.8, 67.3, 67.2, 66.2, 66.1, 60.8, 31.7, 30.4, 30.3, 29.4, 29.12, 29.07, 25.5, 22.3, 21.5, 19.3, 19.22, 19.18, 13.0; ^{31}P NMR (162 MHz, MeOD): δ -10.4 (d, *J* = 16.7 Hz), -13.2 (d, *J* = 16.2 Hz); LRMS (*m/z*): $[M-H]^-$ calcd. for $C_{25}H_{45}NO_{15}P_2$, 660.5; found, 660.5.

Final deprotection with catalytic amount of 0.1% $NaOCH_3$ in CH_3OH and subsequent addition of Amberlyst-15 ion exchange resin to the reaction mixture until around neutral pH (pH 7.0) afforded the GalNAc-PP-Undecyl in quantitative yield. LRMS (*m/z*): $[M-H]^-$ calcd. for $C_{19}H_{38}NO_{12}P_2$, 534.2; found 534.2.

ESI-MS ANALYSIS OF O-UNIT-PP-SUBSTRATES

Following the chemical synthesis of GalNAc-PP-lipid compounds, four glycosyl residues are successively ligated at the non-reducing end by well-characterized glycosyltransferases (WbnH, WbnJ, WbnK and WbnI), allowing for the synthesis of various lipids appended with the same oligosaccharide glycan moiety. ESI-MS was used to monitor the reaction for each of the enzymatic glycosylation steps. The ESI-MS spectra for all of the O-unit-PP-lipid substrates are shown in Fig. S2.

Table S1. The oligonucleotide primers used for cloning and mutagenesis of *waaL*

Primer name	Sequence
$WaaL_{R3}$-NcoI-F	5-ATGC***CCATGG***CTACCTCAACATTATTTTTCTTTCTCGAG-3
$WaaL_{R3}$-HindIII-R	5-ATGC***AAGCTT***TTAGTGGTGGTGGTGGTGGTGGTGGTGGTGCTTGTTTTTCATCGCTAATA-3
$WaaL_{R3}$-R209A-F	5'-TTTCTGATATTAGGAACTTTATCC***GCA***GGGGCCTGGCTTTCC-3'
$WaaL_{R3}$-R209A-R	5'-GGAAAGCCAGGCCCC***TGC***GGATAAAGTTCCTAATATCAGAAA-3'
$WaaL_{R3}$-R268A-F	5-CTAACGTATAAACTTCAACAAACTAATAGTTCTTAT***GCA***TATGCAAATGGTACTCAAGGC-3
$WaaL_{R3}$-R268A-R	5-GCCTTGAGTACCATTTGCATA***TGC***ATAAGAACTATTAGTTTGTTGAAGTTTATACGTTAG-3
$WaaL_{R3}$-H319A-F	5-GACCTTTAGACAATCAATAGGGCCA***GCA***AATTTTGCGCTATTCATCTGGTTTG-3
$WaaL_{R3}$-H319A-R	5-CAAACCAGATGAATAGCGCAAAATT***TGC***TGGCCCTATTGATTGTCTAAAGGTC-3

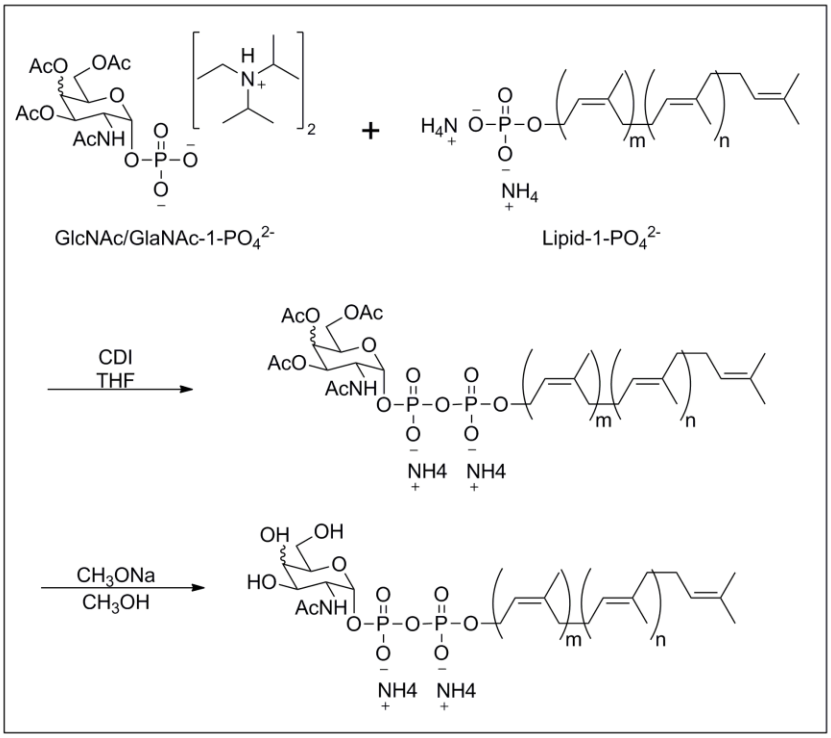

Fig. S1. The general scheme for chemical synthesis of GalNAc/GlcNAc-pp-lipid compounds.

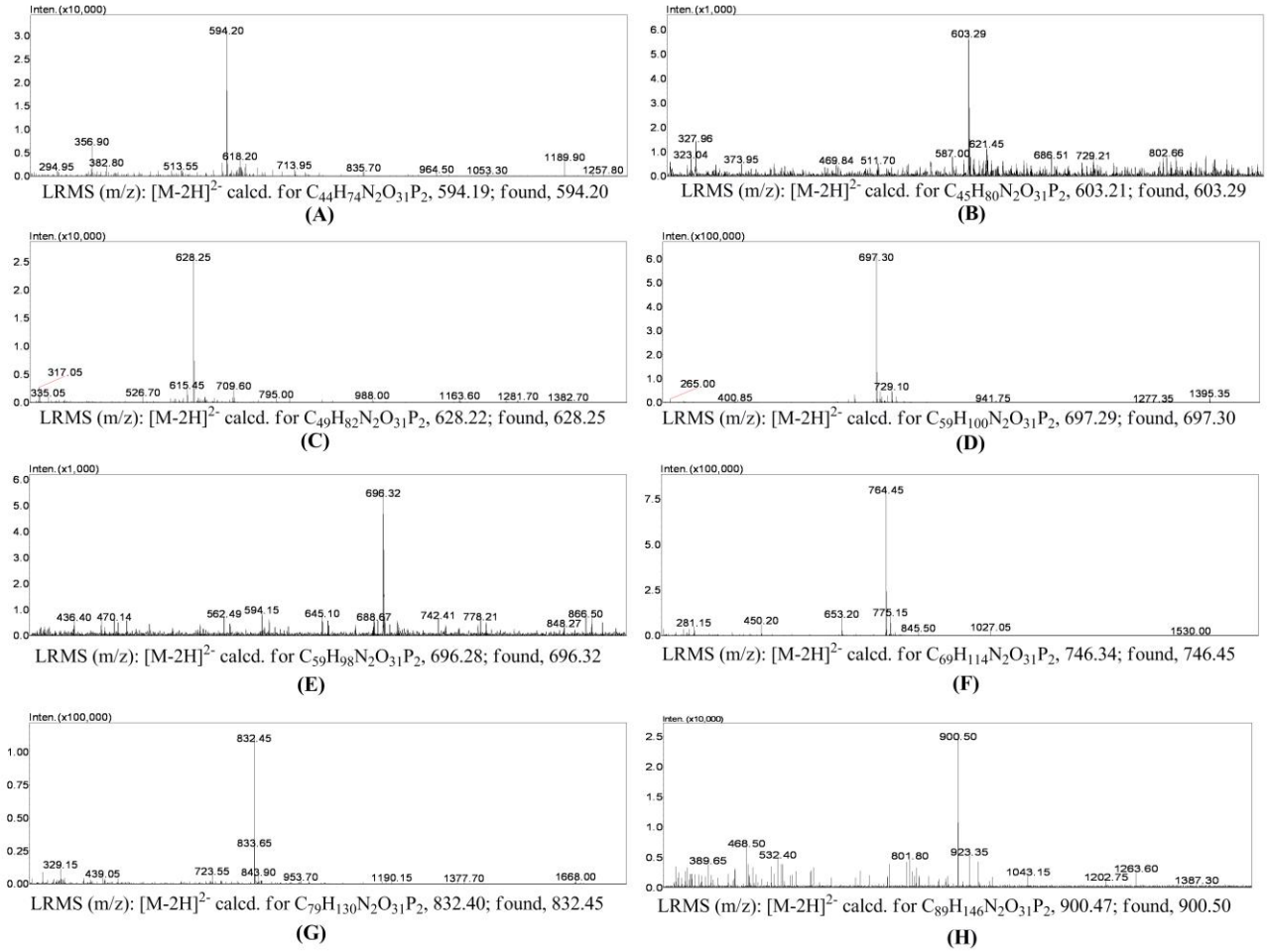

Fig. S2. ESI-MS spectra for all of eight O-unit-PP-lipid substrates. ***A***, O-unit-PP-Geranyl. ***B***, O-unit-PP-Undecyl. ***C***, O-unit-PP-cis-Farnesyl. ***D***, O-unit-PP-MS-Pentaprenyl. ***E***, O-unit-PP-Pentaprenyl. ***F***, O-unit-PP-Heptaprenyl. ***G***, O-unit-PP-Solanesyl. ***H***, O-Unit-PP-Undecaprenyl.

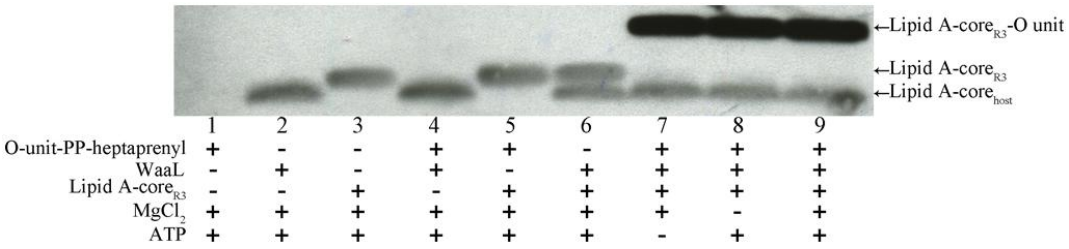

Fig. S3. Western blotting analyses of *in vitro* ligation reaction samples using O-unit-PP-Heptaprenyl donor.

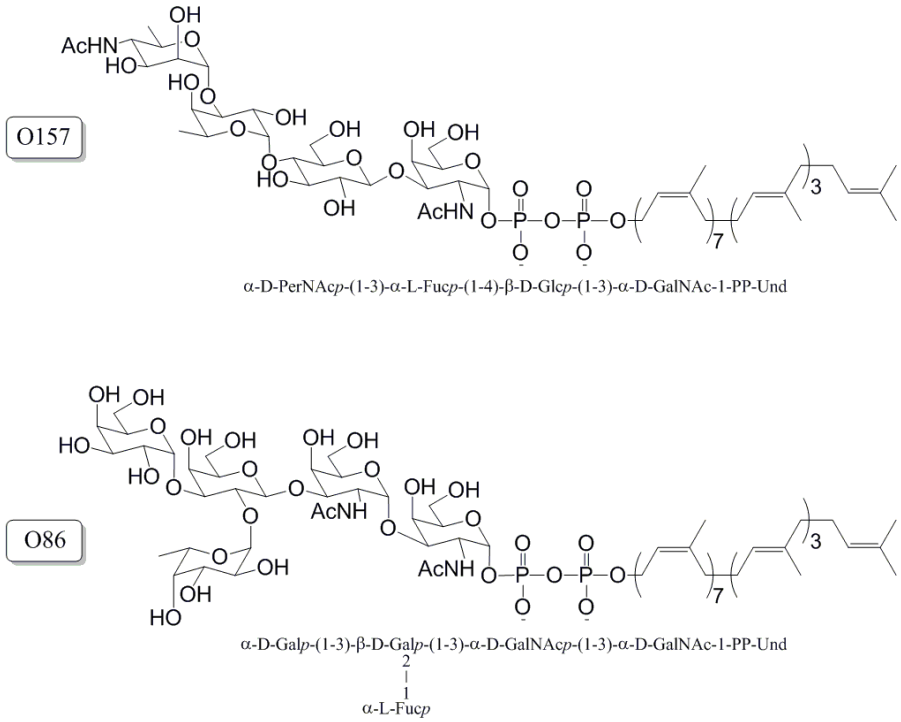

Fig. S4. Structures of O-units of *E. coli* O157 and O86

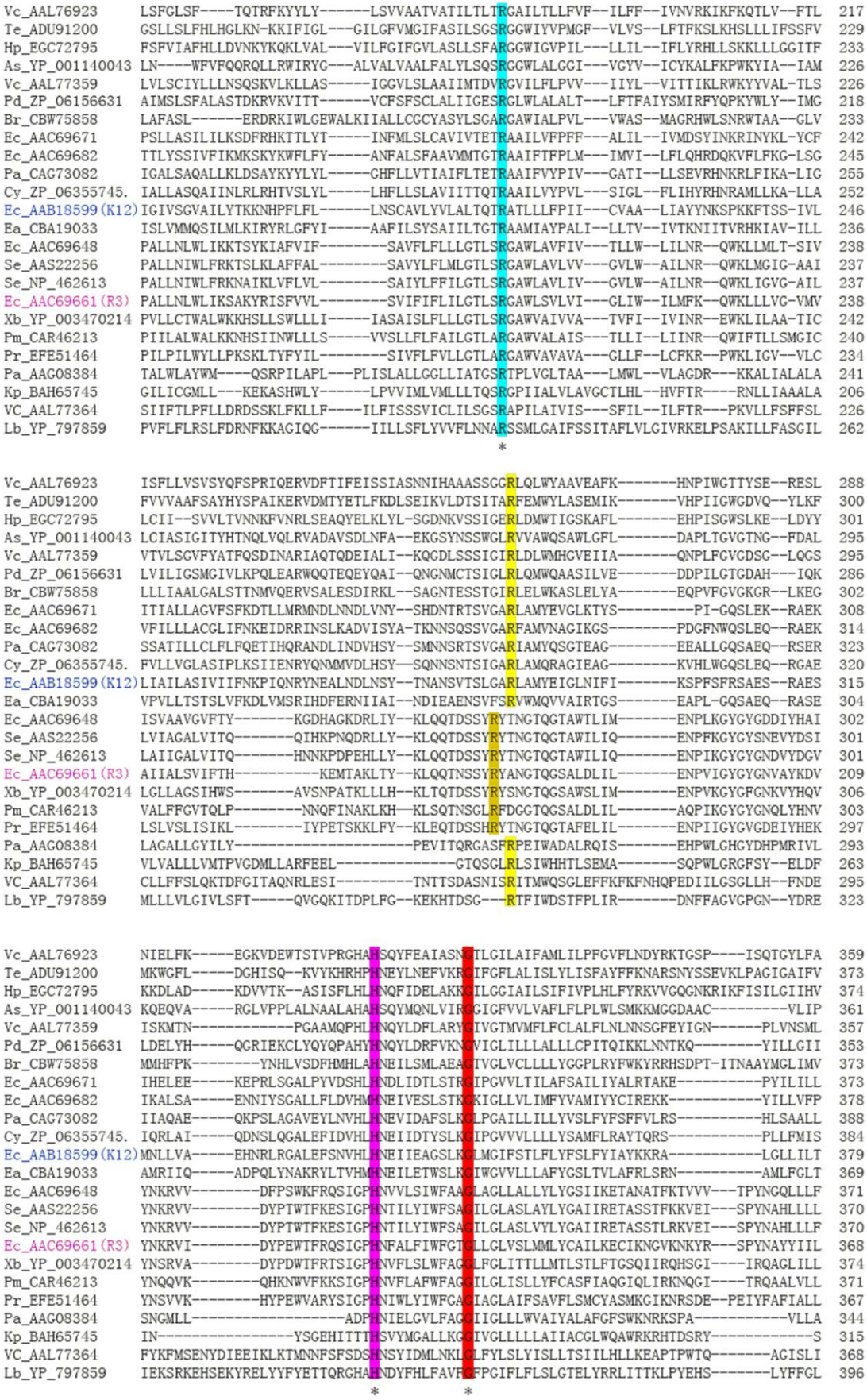
Vc_AAL76923 LSFGLSF----TQTRFKYYLY------LSVVAATVATILTLTRGAILTLLFVF--ILFF--IVNVRKIKFKQTLV--FTL 217
Te_ADU91200 GSLLSLFHLHGLKN-KKIFIGL---GILGFVMGIFASILSGSRGGWIYVPMGF--VLVS--LFTFKSLKHSLLLIFSSFV 229
Hp_EGC72795 FSFVIAFHLLDVNKYKQKLVAL---VILFGIFGVLASLLSFARGGWIGVPIL---LLIL--IFLYRHLLSKKLLLGGITF 233
As_YP_001140043 LN-----WFVFQQRQLLRWIRYG---ALVALVAALFALYLSQSRGGWLALGGI---VGYV--ICYKALFKPWKYIA--IAM 226
Vc_AAL77359 LVLSCIYLLLNSQSKVLKLLAS-----IGGVLSLAAIIMTDVRGVILFLPVV---IIYL--VITTIKLRWKYYVAL-TLS 226
Pd_ZP_06156631 AIMSLSFALASTDKRVKVITT------VCFSFSCLALIIGESRGLWLALALT---LFTFAIYSMIRFYQPKYWLY--IMG 218
Br_CBW75858 LAFASL------ERDRKIWLGEWALKIIALLCGCYASYLSGARGAWIALPVL---VWAS--MAGRHWLSNRWTAA--GLV 233
Ec_AAC69671 PSLLASILILKSDFRHKTTLYT-----INFMLSLCAVIVTETRAAILVFPFF---ALIL--IVMDSYINKRINYKL-YCF 242
Ec_AAC69682 TTLYSSIVFIKMKSKYKWFLFY-----ANFALSFAAVMMTGTRAAIFTFPLM---IMVI--LFLQHRDQKVFLFKG-LSG 245
Pa_CAG73082 IGALSAQALLKLDSAYKYYLYL-----GHFLLVTIAIFLTETRAAIFVYPIV---GATI--LLSEVRHNKRLFIKA-LIG 255
Cy_ZP_06355745. IALLASQAIINLRLRHTVSLYL-----LHFLLSLAVIITTQTRAAILVYPVL---SIGL--FLIHYRHNRAMLLKA-LLA 252
Ec_AAB18599(K12) IGIVSGVAILYTKKNHPFLFL------LNSCAVLYVLALTQTRATLLLFPII---CVAA--LIAYYNKSPKKFTSS-IVL 246
Ea_CBA19033 ISLVMMQSILMLKIRYRLGFYI-----AAFILSYSAIILTGTRAAMIAYPALI--LLTV--IVTKNIITVRHKIAV-ILL 236
Ec_AAC69648 PALLNLWLIKKTSYKIAFVIF--------SAVFLFLLLGTLSRGAWLAVFIV---TLLW--LILNR--QWKLLMLT-SIV 238
Se_AAS22256 PALLNIWLFRKTSLKLAFFAL--------SAVYLFLMLGTLSRGAWLAVLVV---GVLW--AILNR--QWKLMGIG-AAI 237
Se_NP_462613 PALLNIWLFRKNAIKLVFLVL--------SAIYLFFILGTLSRGAWLAVLIV---GVLW--AILNR--QWKLIGVG-AIL 237
Ec_AAC69661(R3) PALLNLWLIKSAKYRISFVVL--------SVIFIFLILGTLSRGAWLSVLVI---GLIW--ILMFK--QWKLLLVG-VMV 238
Xb_YP_003470214 PVLLCTWALWKKHSLLSWLLLI-----IASAISLFLLLGTLSRGAWVAIVVA---TVFI--IVINR--EWKLILAA-TIC 242
Pm_CAR46213 PIILALWALKKNHSIINWLLLS-----VVSLLFLFAILGTLARGAWVALAIS---TLLI--LIINR--QWIFTLLSMGIC 240
Pr_EFE51464 PILPILWYLLPKSKLTYFYIL--------SIVFLFVLLGTLARGAWVAVAVA---GLLF--LCFKR--PWKLIGV--VLC 234
Pa_AAG08384 TALWLAYWM----QSRPILAPL---PLISLALLGGLLIATGSRTPLVGLTAA---LMWL--VLAGDR----KKALIALALA 241
Kp_BAH65745 GILICGMLL---KEKASHWLY------LPVVIMLVMLLLTQSRGPIIALVLAVGCTLHL--HVFTR----RNLLIAAALA 206
VC_AAL77364 SIIFTLPFLLDRDSSKLFKLLF----ILFISSSVICLILSGSRAPILAIVIS---SFIL--ILFTR---PKVLLFSFFSL 226
Lb_YP_797859 PVFLFLRSLFDRNFKKAGIQG------IILLSFLYVVFLNNARSSMLGAIFSSITAFLVLGIVRKELPSAKILLFASGIL 262
Vc_AAL76923 ISFLLVSVSYQFSPRIQERVDFTIFEISSIASNNIHAAASSGGRLQLWYAAVEAFK-------HNPIWGTTYSE--RESL 288
Te_ADU91200 FVVVAAFSAYHYSPAIKERVDMTYETLFKDLSEIKVLDTSITARFEMWYLASEMIK-------VHPIIGWGDVQ--YLKF 300
Hp_EGC72795 LCII--SVVLTVNNKFVNRLSEAQYELKLYL-SGDNKVSSIGERLDMWTIGSKAFL-------EHPISGWSLKE--LDYY 301
As_YP_001140043 LCIASIGITYHTNQLVQLRVADAVSDLNFA--EKGSYNSSWGLRVVAWQSAWLGFL-------DAPLTGVGTNG--FDAL 295
Vc_AAL77359 VTVLSGVFYATFQSDINARIAQTQDEIALI--KQGDLSSSIGIRLDLWMHGVEIIA-------QNPLFGVGDSG--LQGS 295
Pd_ZP_06156631 LVILIGSMGIVLKPQLEARWQQTEQEYQAI--QNGNMCTSIGLRLQMWQAASILVE-------DDPILGTGDAH---IQK 286
Br_CBW75858 LLLIAALGALSTTNMVQERVSALESDIRKL--SAGNTESSTGIRLELWKASLELYA-------EQPVFGVGKGR--LKEG 302
Ec_AAC69671 ITIALLAGVFSFKDTLLMRMNDLNNDLVNY--SHDNTRTSVGARLAMYEVGLKTYS---------PI-GQSLEK--RAEK 308
Ec_AAC69682 VFILLLACGLIFNKEIDRRINSLKADVISYA-TKNNSQSSVGARFAMVNAGIKGS---------PDGFNWQSLEQ--RAEK 314
Pa_CAG73082 SSATILLCLFLFQETIHQRANDLINDVHSY--SMNNSRTSVGARIAMYQSGTEAG---------EEALLGQSAEQ--RSER 323
Cy_ZP_06355745. FVLLVGLASIPLKSIIENRYQNMMVDLHSY---SQNNSNTSIGARLAMQRAGIEAG---------KVHLWGQSLEQ--RGAE 320
Ec_AAB18599(K12) LIAILASIVIIFNKPIQNRYNEALNDLNSY--TNANSVTSLGARLAMYEIGLNIFI-------KSPFSFRSAES--RAES 315
Ea_CBA19033 VPVLLTSTSLVFKDLVMSRIHDFERNIIAI--NDIEAENSVFSRVWMQVVAIRTGS-------EAPL-GQSAEQ--RASE 304
Ec_AAC69648 ISVAAVGVFTY-------KGDHAGKDRLIY--KLQQTDSSYRYTNGTQGTAWTLIM-------ENPLKGYGYGDDIYHAI 302
Se_AAS22256 LVIAGALVITQ-------QIHKPNQDRLLY--KLQQTDSSYRYTNGTQGTAWILIQ-------ENPFKGYGYSNEVYDSI 301
Se_NP_462613 LAIIGALVITQ-------HNNKPDPEHLLY--KLQQTDSSYRYTNGTQGTAWILIQ-------ENPIKGYGYGNDVYDGV 301
Ec_AAC69661(R3) AIIALSVIFTH----------KEMTAKLTY--KLQQTNSSYRYANGTQGSALDLIL-------ENPVIGYGYGNVAYKDV 209
Xb_YP_003470214 LGLLAGSIHWS-------AVSNPATKLLLH--KLTQTDSSYRYSNGTQGSAWSLIM-------ENPVKGYGFGNKVYHQV 306
Pm_CAR46213 VALFFGVTQLP--------NNQFINAKLKH--KLSQTNSGLRFDGGTQGSALDLIL-------AQPIKGYGYGNQLYHNV 303
Pr_EFE51464 LSLVSLISIKL--------IYPETSKKLFY--KLEQTDSSHRYTNGTQGTAFELIL-------ENPIIGYGVGDEIYHEK 297
Pa_AAG08384 LAGALLGYILY---------------------PEVITQRGASFRPEIWADALRQIS-------EHPWLGHGYDHPMRIVL 293
Kp_BAH65745 VLVALLLVMTPVGDMLLARFEEL---------------GTQSGLRLSIWHHTLSEMA-------SQPWLGRGFSY--ELDF 263
VC_AAL77364 CLLFFSLQKTDFGITAQNRLESI---------TNTTSDASNISRITMWQSGLEFFKFKFNHQPEDIILGSGLLH--FNDE 295
Lb_YP_797859 MLLLVLGIVLSFT-----QVGQKITDPLFG--KEKHTDSG---RTFIWDSTFPLIR-------DNFFAGVGPGN--YDRE 323
Vc_AAL76923 NIELFK-----EGKVDEWTSTVPRGHAHSQYFEAIASNGTLGILAIFAMLILPFGVFLNDYRKTGSP-----ISQTGYLFA 359
Te_ADU91200 MKWGFL-----DGHISQ--KVYKHRHPHNEYLNEFVKRGIFGFLALISLYLISFAYFFKNARSNYSSEVKLPAGIGAIFV 373
Hp_EGC72795 KKDLAD-----KDVVTK--ASISFLHLHNQFIDELAKKGILGGIAILSIFIVPLHLFYRKVVGQGNKRIKFISILGIIHV 374
As_YP_001140043 KQEQVA-----RGLVPPLALNAALAHAHSQYMQNLVIRGGIGFVVLVAFLFLPLWLSMKKMGGDAAC---------VLIP 361
Vc_AAL77359 ISKMTN------------PGAAMQPHLHNQYLDFLARYGIVGTMVMFLFCLALFLNLNNSGFEYIGN------PLVNSML 357
Pd_ZP_06156631 LDELYH-----QGRIEKCLYQYQPAHYHNQYLDRFVKNGVIGLILLLALLLCPITQIKKLNNTKQ--------YILLGII 353
Br_CBW75858 MMHFPK--------YNHLVSDFHMHLAHNEILSMLAEAGTVGLVCLLLLYGGPLRYFWKYRRHSDPT-ITNAAYMGLIMV 373
Ec_AAC69671 IHELEE-----KEPRLSGALPYVDSHLHNDLIDTLSTRGIPGVVLTILAFSAILIYALRTAKE----------PYILILL 373
Ec_AAC69682 IKALSA-----ENNIYSGALLFLDVHMHNEIVESLSTKGKIGLLVLIMFYVAMIYYCIREKK----------YILLVFP 378
Pa_CAG73082 IIAQAE-----QKPSLAGAVEYLNVHLHNEVIDAFSLKGLPGAILLILLYVSLFYFSFFVLRS----------HLSAALL 388
Cy_ZP_06355745. IQRLAI-----QDNSLQGALEFIDVHLHNEIIDTYSLKGIPGVVVLLLLYSAMFLRAYTQRS----------PLLFMIS 384
Ec_AAB18599(K12) MNLLVA-----EHNRLRGALEFSNVHLHNEIIEAGSLKGLMGIFSTLFLYFSLFYIAYKKRA----------LGLLILT 379
Ea_CBA19033 AMRIIQ-----ADPQLYNAKRYLTVHMHNEILETWSLKGIWGVVLLLAFYGSLTVLAFRLSRN----------AMLFGLT 369
Ec_AAC69648 YNKRVV--------DFPSWKFRQSIGPHNVVLSIWFAAGLAGLLALLYLYGSIIKETANATFKTVVV---TPYNGQLLLF 371
Se_AAS22256 YNKRVV--------DYPTWTFKESIGPHNTILYIWFSAGILGLASLAYLYGAIIRETASSTFKKVEI---SPYNAHLLLL 370
Se_NP_462613 YNKRVV--------DYPTWTFKESIGPHNTILYIWFSAGILGLASLVYLYGAIIRETASSTLRKVEI---SPYNAHLLLF 370
Ec_AAC69661(R3) YNKRVI--------DYPEWTFRQSIGPHNFALFIWFGTGLLGLVSLMMLYCAILKECIKNGVKNKYR---SPYNAYYIIL 368
Xb_YP_003470214 YNSRVA--------DYPDWTFRTSIGPHNVFLSLWFAGGLFGLITTLLMTLSTLFTGSQIIRQHSGI-----IRQAGLILL 374
Pm_CAR46213 YNQQVK--------QHKNWVFKKSIGPHNVFLAFWFAGGILGLISLLYFCASFIAQGIQLIRKNQGI-----TRQAALVLL 371
Pr_EFE51464 YNSVVK--------HYPEWVARYSIGPHNIWLYIWFGAGIAGLAIFSAVFLSMCYASMKGIKNRSDE--PEIYFAFIALL 367
Pa_AAG08384 SNGMLL-------------------ADPHNIELGVLFAGGIIGLLLWVAIYALAFGFSWKNRKSPA----------VLLA 344
Kp_BAH65745 IN----------------YSGEHITTTHSVYMGALLKGGIVGLLLLLAIIACGLWQAWRKRHTDSRY----------S 315
VC_AAL77364 FYKFMSENYDIEEIKLKTMNNFSFSDSHNSYIDMLNKLGLFYLSLYISLLTSIILHLLKEAPTPWTQ-------AGISLI 368
Lb_YP_797859 IEKSRKEHSEKYRELYYFYETTQRGHAHNDYFHLFAVFGFPGIFLFLSLGTELYRRLITTKLPYEHS-------LYFFGL 396

Fig. S5. Multiple sequence alignment analysis of WaaL homologues.