Supporting Information

Characterization of PglJ, a Glycosyltransferase in the Campylobacter concisus N-linked Protein Glycosylation Pathway that Expands Glycan Diversity

Christine A. Arbour,[a,b] Nemanja Vuksanovic,[c] Hannah M. Bernstein,[a] Karen N. Allen,[c] Barbara Imperiali[a,b]*

[a] Department of Biology, Massachusetts Institute of Technology, 31 Ames St, Cambridge, MA 02139, USA

[b] Department of Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA

[c] Department of Chemistry, Boston University, 590 Commonwealth Ave, Boston, MA 02215, USA

***Corresponding Author**

Corresponding author:

Barbara Imperiali

Tel.: +1 617 253 1838

Email: imper@mit.edu

Table of Contents

General Information **2**

UDP-GlcNAcA chemical oxidation **3**

UDP-GalNAcA chemical oxidation **7**

Expression and purification of glycosyltransferase enzymes **13**

Und-PP-diNAcBac-GalNAc enzymatic synthesis **18**

UDP-Glo biochemical assays to quantify Und-PP-diNAcBac-GalNAc **19**

UDP Glo coupled-enzyme biochemical assays **24**

Expression and purification of C. concisus strain 13826 TviB family dehydrogenase (Cc DH) .. **25**

Campylobacter concisus dehydrogenase (13826) assays **26**

References **28**

General Information

Unless otherwise specified, all commercially available reagents were used without further purification. Normal phase chromatographic purifications were conducted using SiliaFlash Irregular Silica Gel P60, 40 - 63 μm, 60 Å (SiliCycle). Thin-layer chromatography (TLC) was performed on SiliaPlate glass-backed, silica gel TLC plates (250 μm, F254, SiliCycle TLG-R10014B-323). ^{1}H and ^{13}C NMR spectra were recorded on a Bruker 600 MHz spectrometer. ^{31}P NMR spectra were recorded on a Bruker Avance Neo spectrometer operating at 500.34 MHz. All ^{13}C spectra are recorded with complete proton decoupling. Chemical shifts are reported in ppm from tetramethylsilane (TMS) with the solvent resonance as the internal standard. For samples in D_2O, the spectrometer uses the absolute deuterium (D_2O) lock frequency for all multinuclear chemical shift calibrations. Additionally, yields refer to chromatographically and spectroscopically pure products. High-resolution mass spectral analyses were obtained on an Agilent 6545 QTOF mass spectrometer with a Jet Stream ESI source coupled to an Agilent Infinity 1260 LC system.

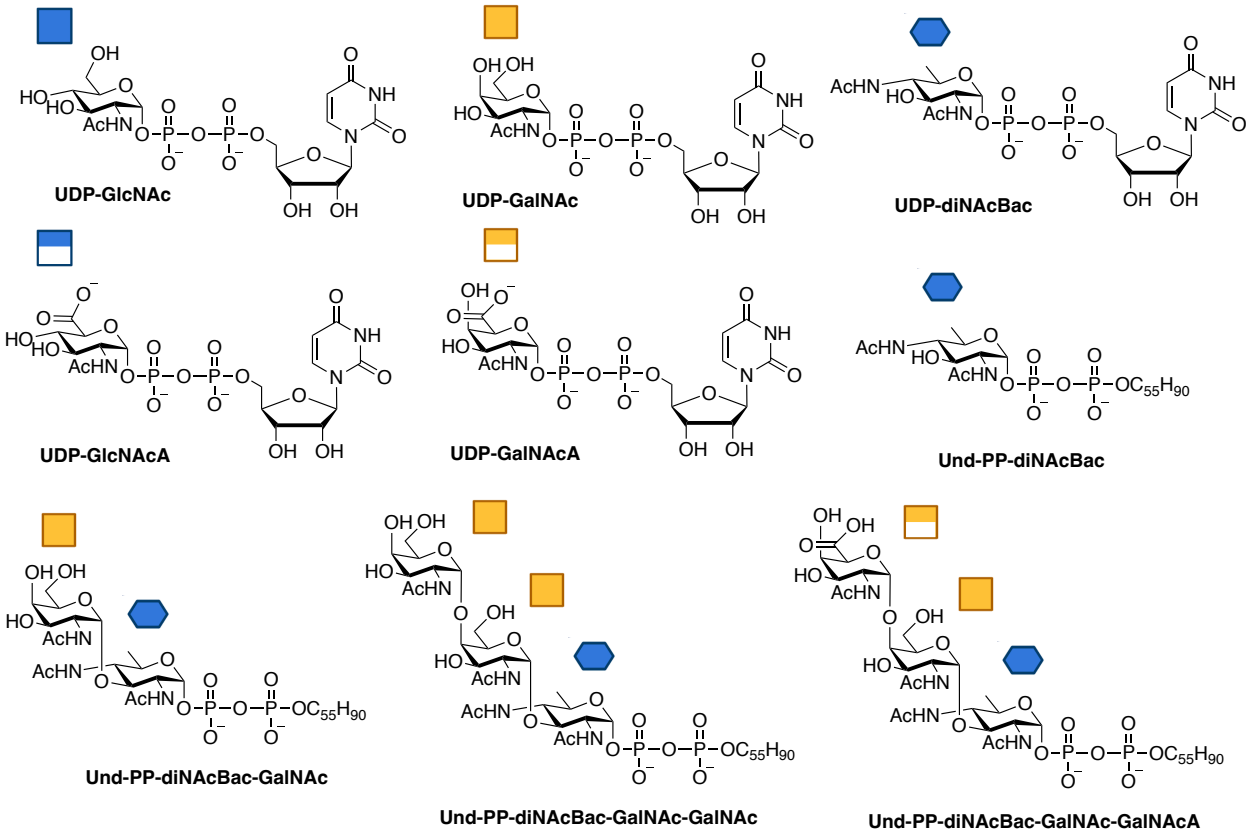
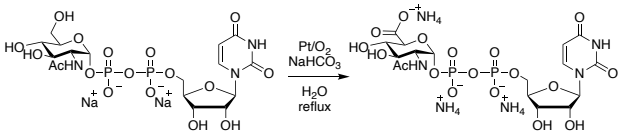


Figure S1. Acceptor and donor substrates used in this study with the associated monosaccharide symbols.[1]



UDP-GlcNAcA chemical oxidation

(2*S*,3*S*,4*R*,5*R*,6*R*)-5-acetamido-6-(((((((2*R*,3*S*,4*R*,5*R*)-5-(2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methoxy)(hydroxy)phosphoryl)oxy)(hydroxy)phosphoryl)oxy)-3,4-dihydroxytetrahydro-2*H*-pyran-2-carboxylic acid. UDP-GlcNAcA was prepared via oxidation of UDP-GlcNAc as previously described.[2] Adams' catalyst (Aldrich; 5 mg) was hydrogenated in water (1 mL) under atmospheric pressure for 2 h. The brown slurry of catalyst turned into a black clumpy suspension once hydrogenated. The apparatus was evacuated and released under nitrogen 3 times. The freshly prepared platinum in water was added to a solution

of UDP-GlcNAc (20 mg, 0.31 mmol) and $NaHCO_3$ (9.1 mg, 0.108 mmol) in water (1.12 mL, 0.0275 M rel. to UDP-sugar). The vigorously stirred mixture was heated to 100 °C and oxygen was continually passed into the mixture through a 22 G x 4" hypodermic needle (Air-Tite) submerged into the refluxing solution. After 5 h, 21 h, and 31 h a new batch of catalyst, prepared by hydrogenation of Adams' catalyst (5 mg) in water (1 mL), was added. After 48 h, the catalyst was removed by centrifugation to pellet the catalyst, followed by filtration of the supernatant through a 0.2 µm filter (VWR # 28145-475). The solution was purified using anion exchange FPLC using tandem Cytiva HiTrap™ 5 mL Q HP columns (2x) with mobile phases of A: Milli-Q H_2O and B: 0.5 M NH_4HCO_3 with a gradient of 25% B over 10 mL, then 25-100% B over 140 mL. Fractions containing the title compound were pooled and lyophilized to provide UDP-GlcNAcA (2.5 mg, 11% yield) as a white solid. The desired UDP-GlcNAcA was characterized by HRMS and ^{1}H, ^{13}C, and ^{31}P NMR spectroscopy, and the resulting data were found to match previously published values.[2, 3]

^{1}H NMR (600 MHz, D_2O) δ 7.86 (d, *J* = 8.1 Hz, 1H), 5.88 (dd, *J* = 10.9, 6.2 Hz, 2H), 5.44 (dd, *J* = 7.7, 3.3 Hz, 1H), 4.27 (t, *J* = 3.4 Hz, 2H), 4.19 (t, *J* = 3.0 Hz, 1H), 4.14 (ddd, *J* = 14.0, 4.6, 2.2 Hz, 1H), 4.11 – 4.05 (m, 2H), 3.93 (dt, *J* = 10.6, 3.0 Hz, 1H), 3.73 (dd, *J* = 10.5, 9.1 Hz, 1H), 3.50 (t, *J* = 9.6 Hz, 1H), 1.98 (s, 3H).

^{13}C NMR (151 MHz, D_2O) δ 176.0 (1C, s, *C*OO^-), 174.8 (1C, s, CH_3*C*O•NH), 166.6 (1C, s, C-4), 151.8 (1C, s, C-2), 142.1 (1C, s, C-6), 102.6 (1C, s, C-5), 94.19 (1C, dd, $J_{C\text{-}1'',Pb}$ = 6.0 Hz, C-1"), 88.3 (1C, s, C-1'), 83.2 (1C, dd, $J_{C\text{-}4',Pa}$ = 8.7 Hz, C-4'), 73.8 (1C, s, C-2'), 73.1 (1C, s, C-5"), 72.1 (1C, s, C-4"), 70.6 (1C, s, C-3"), 69.6 (1C, s, C-3'), 64.9 (1C, td, $J_{C\text{-}5',Pa}$ = 5.4 Hz, C-5'), 53.4 (1C, dd, $J_{C\text{-}2'',Pb}$ = 8.2 Hz, C-2"), 22.0 (1C, s, *C*H_3CO•NH).

^{31}P NMR (203 MHz, D_2O) δ -11.32 (P_β, d, $J_{P\alpha,P\beta}$ = 19.8 Hz), -13.06 (Pα, d, $J_{P\alpha,P\beta}$ = 20.2 Hz).

HRMS (ESI-) *m/z* calcd for $C_{17}H_{24}N_3O_{18}P_2$ $[M\text{-}H]^-$ 620.0536, found 620.0536.

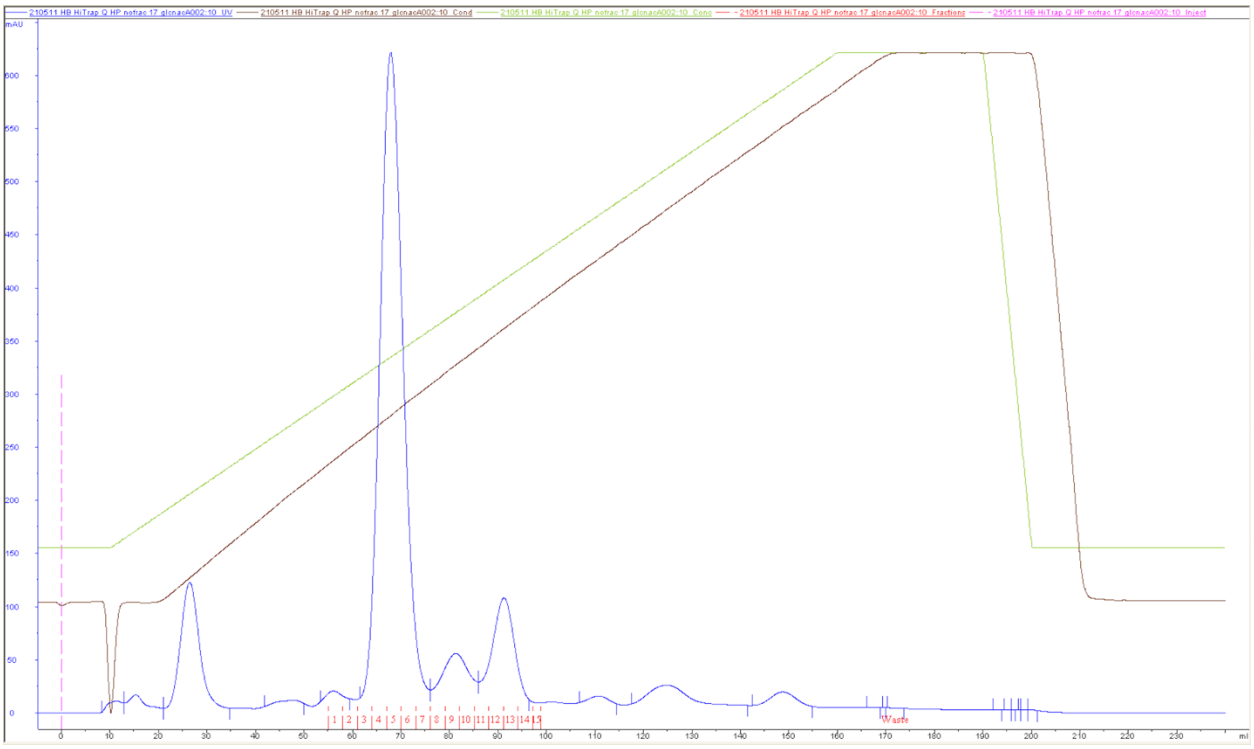


Figure S2. FPLC spectrum of the crude UDP-GlcNAcA reaction monitored at 260 nm. UDP-GlcNAcA elutes from 62-75 mL.

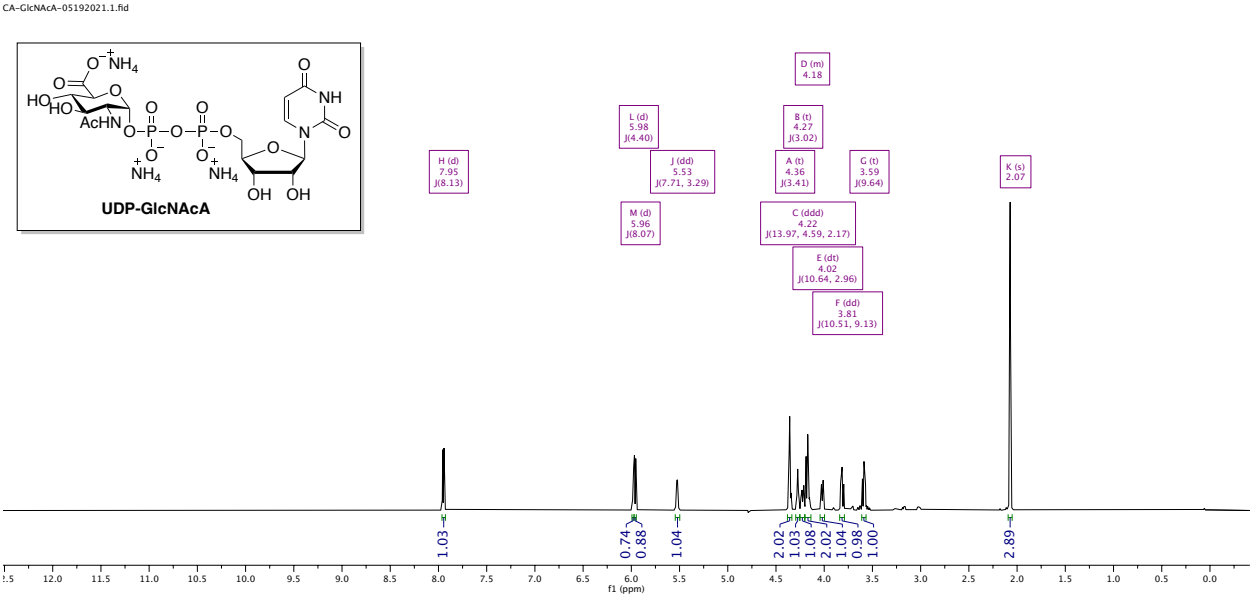


Figure S3. ^{1}H NMR spectrum (600 MHz, D_2O) of UDP-GlcNAcA. The spectrum was acquired at 25 ºC with suppression of the HOD signal at 4.80 ppm.

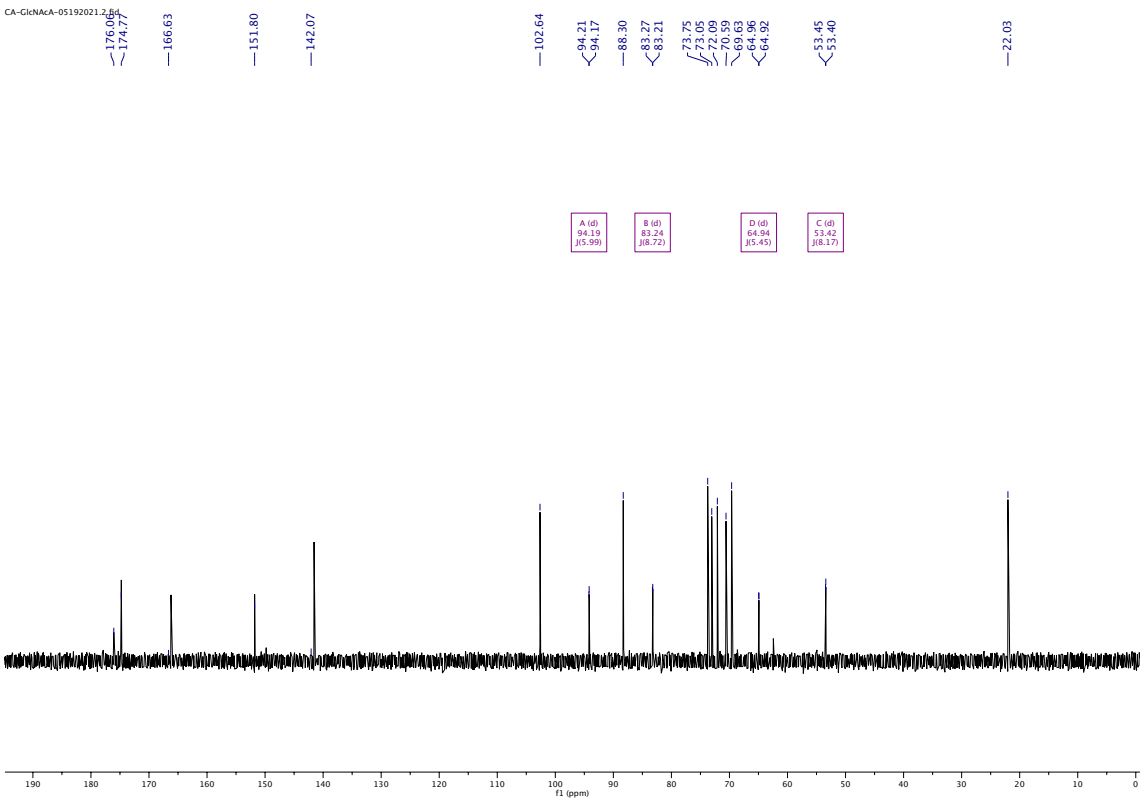


Figure S4. ^{13}C NMR spectrum (151 MHz, D_2O) of UDP-GlcNAcA.

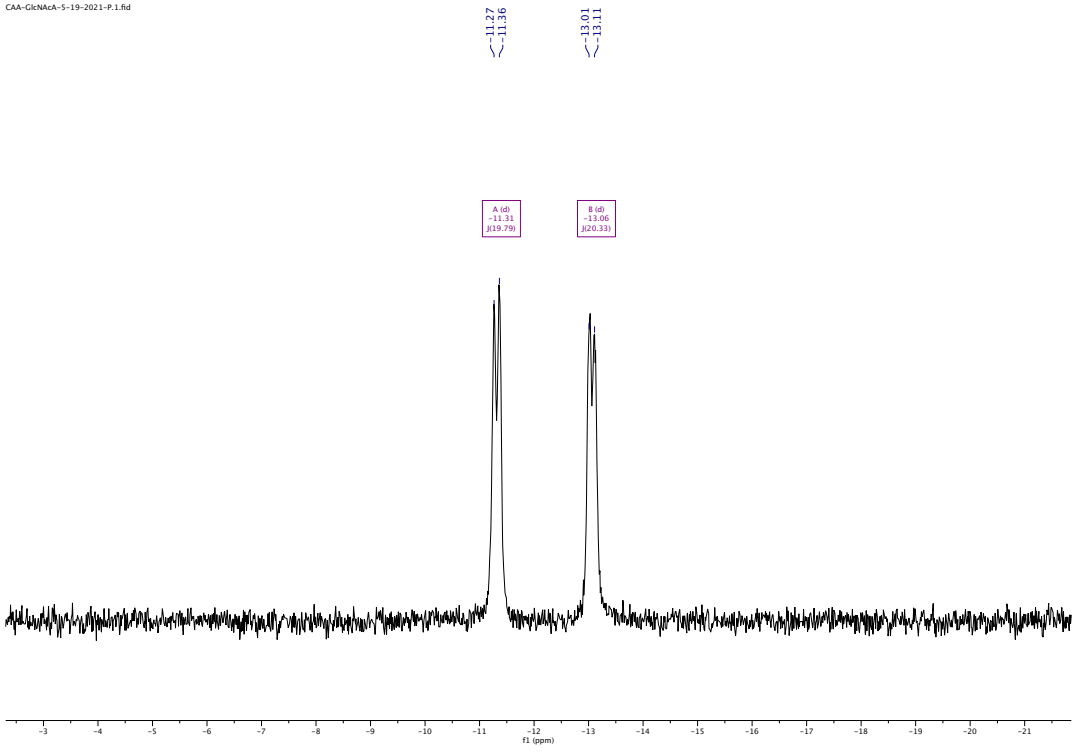
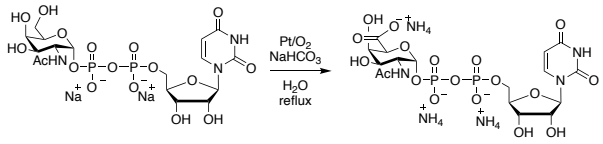


Figure S5. ^{31}P NMR spectrum (203 MHz, D_2O) of UDP-GlcNAcA.

Table S1. Comparison of NMR chemical shifts and *J* couplings of UDP-GlcNAcA with literature values

	UDP-GlcNAcA			UDP-GlcNAcA Lit. (Field *Carbohydr. Res* . 2007)			UDP-GlcNAcA Lit. (Imperiali *Biochem.* 2009)			UDP-GlcNAcA Lit. (Lam *J. Biol. Chem.* **2008**)		
Moiety	**δ (ppm)**	***J***	**Hz**	**δ (ppm)**	***J***	**Hz**	**δ (ppm)**	***J***	**Hz**	**δ (ppm)**	***J***	**Hz**
Uracil												
H-5	5.96	$J_{5,6}$	8.1	5.96	$J_{5,6}$	8.2	5.95	$J_{5,6}$	8.1	5.97		
H-6	7.95			7.95			7.94			7.95		
Ribose												
H-1'	5.98	$J_{1',2'}$	4.4	5.98	$J_{1',2'}$	4.4	5.97	$J_{1',2'}$	4.4	5.97		
H-2'	4.36			4.38–4.34			4.36			4.37		
H-3'	4.36			4.38–4.34			4.36			4.36		
H-4'	4.27			4.29–4.27			4.29			4.28		
H-5'a	4.22			4.24–4.20			4.23			4.23		
H-5'b	4.19–4.14			4.18–4.13			4.17	$J_{5'a,5'b}$	11.7	4.18		
Pyranose												
H-1''	5.53	$J_{1'',P}$	7.7	5.52	$J_{1'',P}$	7.6	5.57	$J_{1'',P}$	7.4	5.54	$J_{1'',P(\beta)}$	7.6
		$J_{1'',2''}$	3.3		$J_{1'',2''}$	3.2		$J_{1'',2''}$	3.2		$J_{1'',2''}$	3.3
H-2''	4.02	$J_{2'',P}$	3.0	4.02	$J_{2'',P}$	3.0	4.32			4.02		
		$J_{2'',3''}$	10.6		$J_{2'',3''}$	10.6		$J_{2'',3''}$	11		$J_{2'',3''}$	10.5
H-3''	3.81	$J_{3'',4''}$	10.5	3.81	$J_{3'',4''}$	9.4	3.8	$J_{3'',4''}$	10.1	3.82	$J_{3'',4''}$	9.9
H-4''	3.59	$J_{4'',5''}$	9.6	3.58	$J_{4'',5''}$	10.5	3.57	$J_{4'',5''}$	10	3.59	$J_{4'',5''}$	9.6
H-5''	4.19–4.14			4.16			4.19			4.17		
acetyl	2.07			2.07			2.09			2.08		



UDP-GalNAcA chemical oxidation

(2*S*,3*R*,4*R*,5*R*,6*R*)-5-acetamido-6-(((((((2*R*,3*S*,4*R*,5*R*)-5-(2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methoxy)(hydroxy)phosphoryl)oxy)(hydroxy)phosphoryl)oxy)-3,4-dihydroxytetrahydro-2*H*-pyran-2-carboxylic acid. UDP-GalNAcA was prepared via oxidation of UDP-GalNAc with a modified procedure that was used for UDP-GlcNAcA.**Error! Bookmark not defined.**,**Error! Bookmark not defined.** Adams' catalyst (Aldrich; 5 mg) was hydrogenated in water (1 mL) under atmospheric pressure for 2 h. The brown slurry of catalyst turned into a black clumpy suspension once hydrogenated. The apparatus was evacuated and released under nitrogen 3 times. The freshly prepared platinum in water was added to a solution of UDP-GalNAc (5 mg, 8.2 μmol) and $NaHCO_3$ (7.6 mg, 0.09 mmol) in water (2.5 mL). The vigorously stirred mixture was heated to 100 °C and oxygen was continually passed into the mixture through a 22 G x 4" hypodermic needle (Air-Tite) submerged into the refluxing solution. After 5 h, 21 h, and 31 h a new batch of catalyst, prepared by hydrogenation of

Adams' catalyst (5 mg) in water (1 mL), was added. After 48 h, the catalyst was removed by centrifugation to pellet the catalyst, followed by filtration of the supernatant through a 0.2 μm filter (VWR # 28145-475). The solution was purified using anion exchange FPLC using tandem Cytiva HiTrap^{TM} 5 mL Q HP columns (2x) with mobile phases of A: Milli-Q H_2O and B: 0.5 M NH_4HCO_3 with a gradient of 25% B over 10 mL, then 25-100% B over 140 mL. Fractions containing the title compound were pooled and lyophilized to provide UDP-GalNAcA (0.8 mg, 16% yield) as a white solid. The desired UDP-GalNAcA was characterized by HRMS and ^{1}H, ^{13}C, and ^{31}P NMR spectroscopy, and the resulting data were found to match previously published values.[3] The optimal length of the reaction that minimizes the amount of degradation is 12 h, as visualized by FPLC.

^{1}H NMR (600 MHz, D_2O) δ 7.95 (d, *J* = 8.2 Hz, 1H), 5.98 (d, *J* = 4.5 Hz, 1H), 5.96 (d, *J* = 8.1 Hz, 1H), 5.58 (dd, *J* = 7.3, 3.4 Hz, 1H), 4.49 (d, *J* = 1.5 Hz, 1H), 4.36 (q, *J* = 4.1, 3.1 Hz, 2H), 4.32 (d, *J* = 3.1 Hz, 1H), 4.29 – 4.24 (m, 2H), 4.21 (ddd, *J* = 11.9, 4.6, 2.6 Hz, 1H), 4.15 (ddt, *J* = 11.7, 5.4, 3.2 Hz, 1H), 4.02 (dd, *J* = 11.0, 3.2 Hz, 1H), 2.08 (s, 3H).
^{13}C NMR (151 MHz, D_2O) δ 175.2 (1C, s, *C*OO^-), 175.0 (1C, s, CH_3*C*O•NH), 166.3 (1C, s, C-4), 151.8 (1C, s, C-2), 141.6 (1C, s, C-6), 102.6 (1C, s, C-5), 94.5 (1C, d, $J_{C\text{-}1'',P\beta}$ = 6.5 Hz, C-1"), 88.3 (1C, s, C-1'), 83.2 (1C, d, $J_{C\text{-}4',P\alpha}$ = 9.3 Hz, C-4'), 73.7 (1C, s, C-2'), 73.0 (1C, s, C-5"), 69.8 (1C, s, C-3"), 69.6 (1C, s, C-3'), 67.6 (1C, s, C-4"), 64.9 (1C, d, $J_{C\text{-}5',P\alpha}$ = 5.4 Hz, C-5'), 49.4 (1C, d, $J_{C\text{-}2'',P\beta}$ = 8.7 Hz, C-2"), 22.1 (1C, s, *C*H_3CO•NH).
^{31}P NMR (203 MHz, D_2O) δ -11.34 (P_β, d, $J_{P\alpha,P\beta}$ = 20.0 Hz), -12.78 (P_α, dd, $J_{P\alpha,P\beta}$ = 20.0, 6.5 Hz).
HRMS (ESI-) *m/z* calcd for $C_{17}H_{24}N_3O_{18}P_2$ $[M\text{-}H]^-$ 620.0536, found 620.0540.

Figure S6. FPLC spectrum of the crude UDP-GalNAcA reaction monitored at 260 nm after 6 h. UDP-GalNAcA elutes from 60-70 mL.

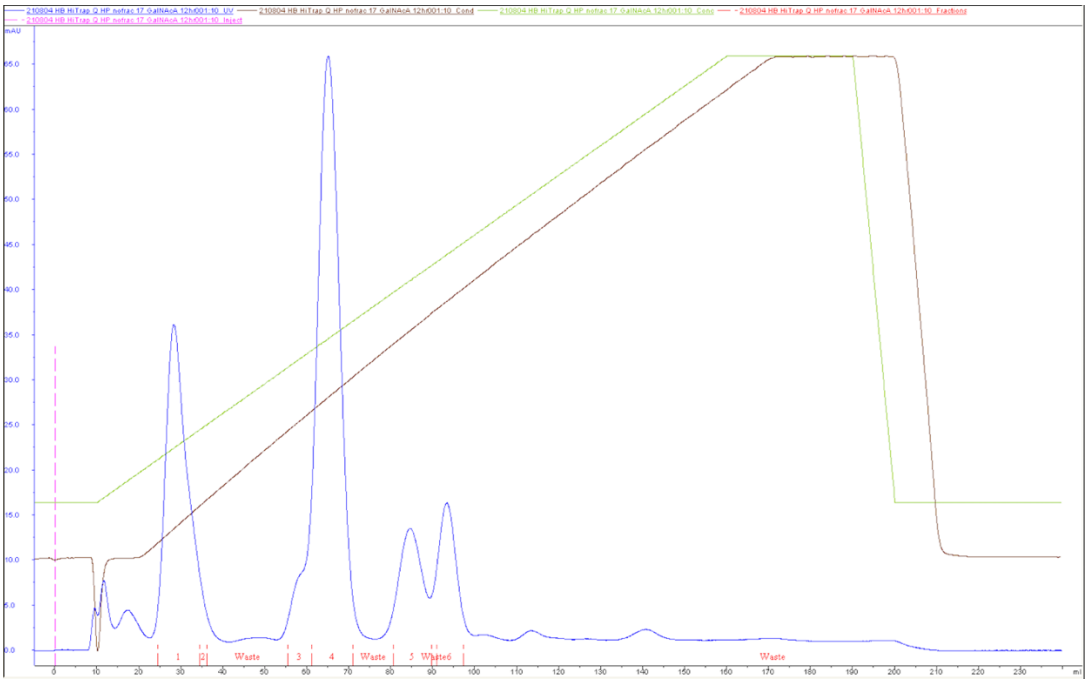


Figure S7. FPLC spectrum of the crude UDP-GalNAcA reaction monitored at 260 nm after 12 h. UDP-GalNAcA elutes from 60-70 mL.

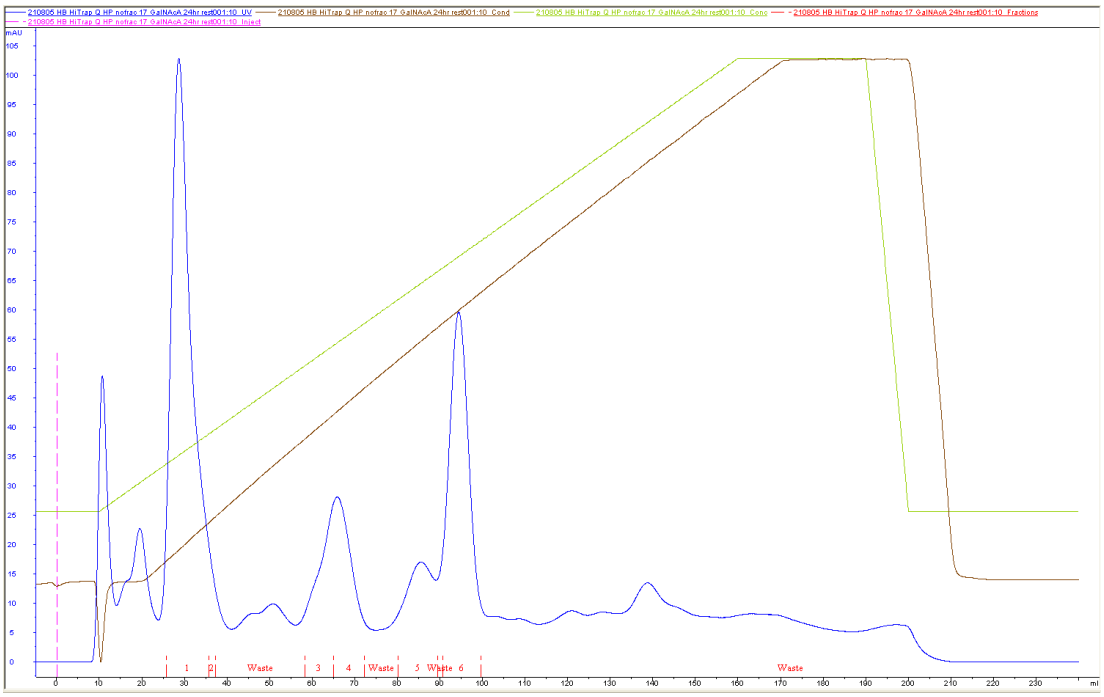


Figure S8. FPLC spectrum of the crude UDP-GalNAcA reaction monitored at 260 nm after 24 h. UDP-GalNAcA elutes from 60-70 mL.

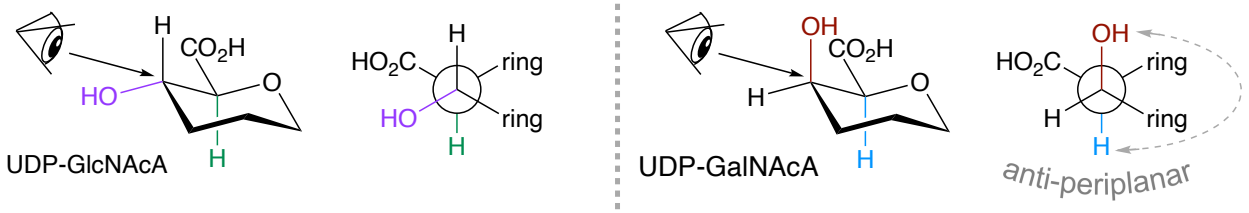


Figure S9. Simplified chair conformations and Newman projections for GlcNAcA and GalNAcA.

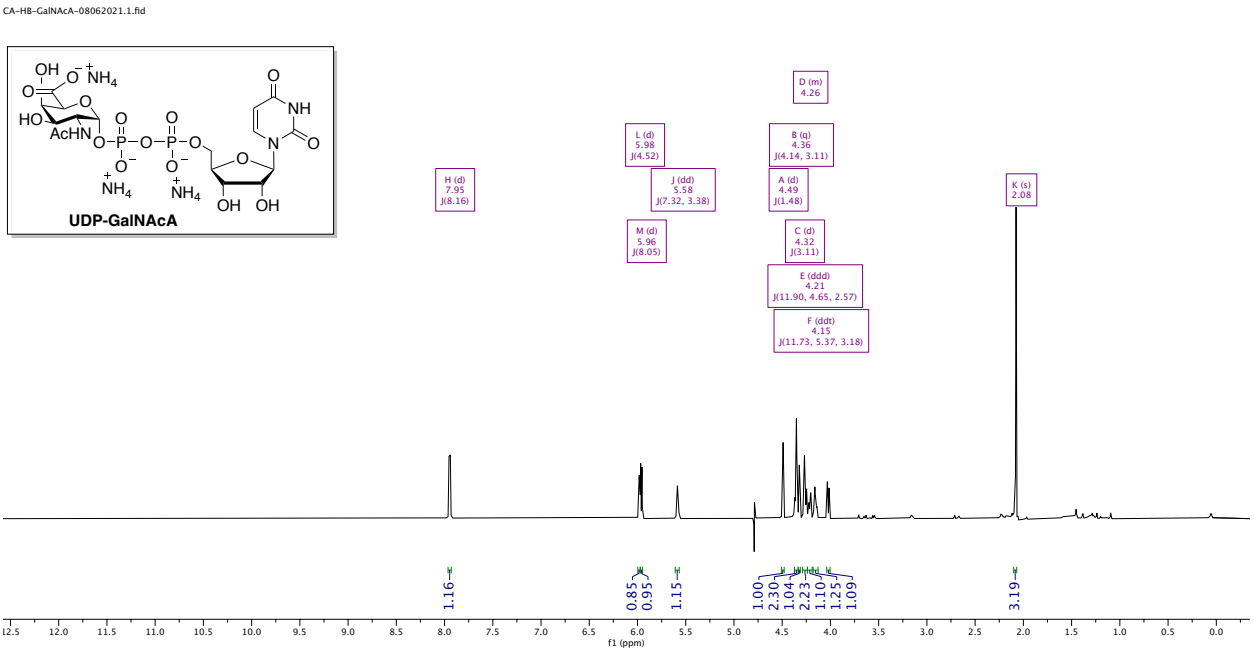


Figure S10. ^{1}H NMR spectrum (600 MHz, D_2O) of UDP-GalNAcA. The spectrum was acquired at 25 ºC with suppression of the HOD signal at 4.80 ppm.

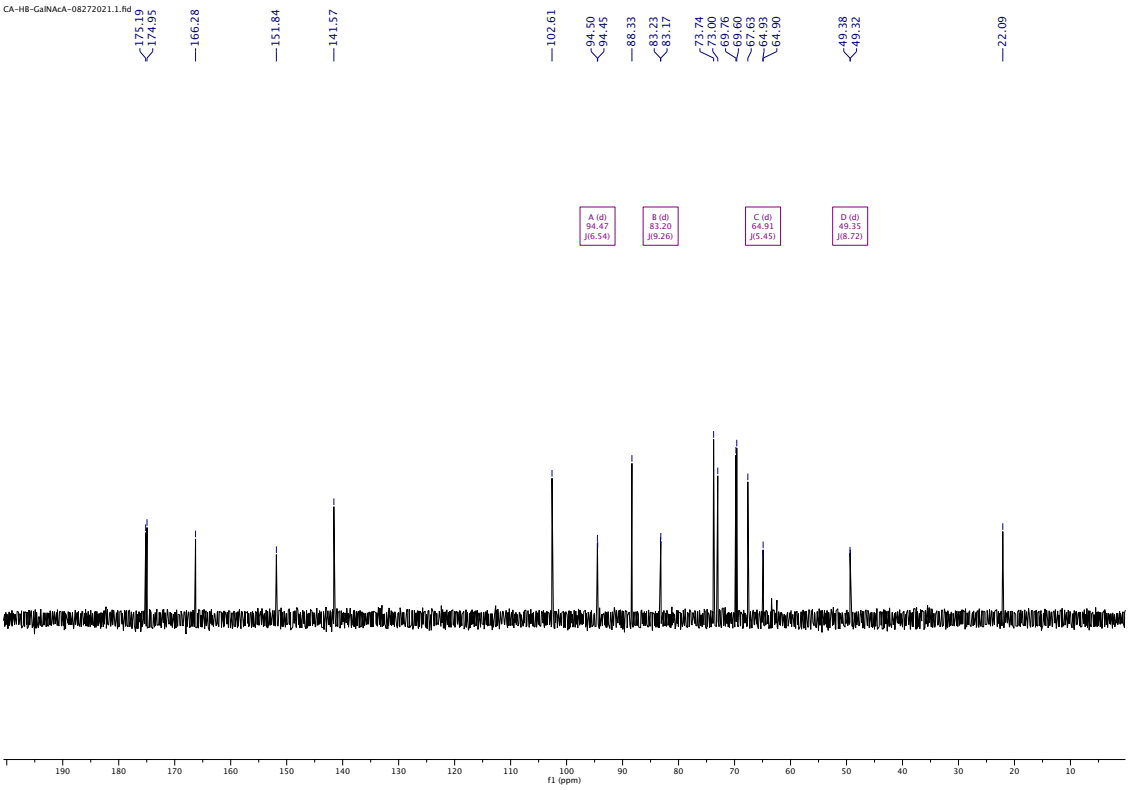


Figure S11. ^{13}C NMR spectrum (151 MHz, D_2O) of UDP-GalNAcA.

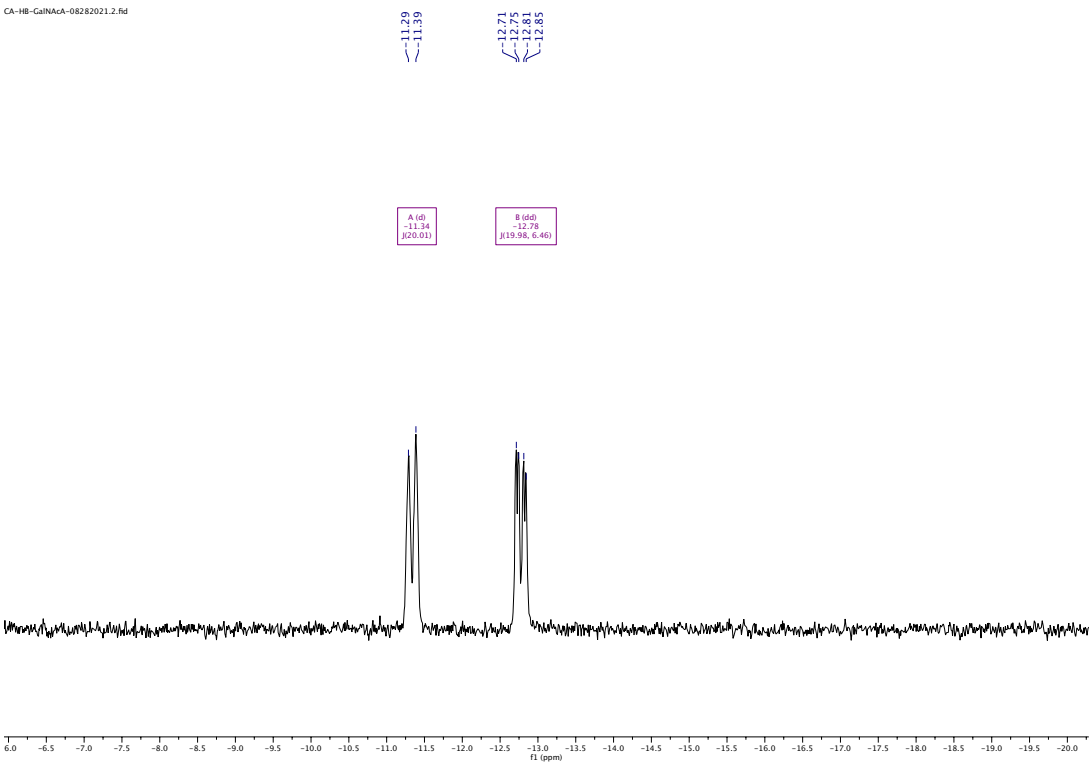


Figure S12. ^{31}P NMR spectrum (203 MHz, D_2O) of UDP-GalNAcA.

Table S2. Comparison of NMR chemical shifts and *J* couplings of UDP-GalNAcA with literature values.

	UDP-GalNAcA Lit. (Lam *J. Biol. Chem.* **2008**)			UDP-GalNAcA		
Moiety	**δ (ppm)**	***J***	**Hz**	**δ (ppm)**	***J***	**Hz**
Uracil						
H-5	5.97			5.96	$J_{5,6}$	8.1
H-6	7.95			7.95		
Ribose						
H-1'	5.97			5.98	$J_{1',2'}$	4.5
H-2'	4.37			4.36	$J_{2',3'}$	4.1, 3.1
H-3'	4.36			4.36	$J_{3',4'}$	4.1, 3.1
H-4'	4.28			4.26		
H-5a'	4.23			4.21	$J_{5'a,5'b}$	11.9, 4.7, 2.6
H-5b'	4.18			4.18		11.7, 5.4, 3.2
Pyranose						
H-1''	5.61	$J_{1'',P(\beta)}$	7.1	5.58	$J_{1'',P}$	7.3
		$J_{1'',2''}$	3.3		$J_{1'',2''}$	3.4
H-2''	4.26	$J_{2'',3''}$	11.1	4.26	$J_{2'',3''}$	11.0
H-3''	4.03	$J_{3'',4''}$	3.1	4.02	$J_{3'',4''}$	3.2
H-4''	4.33	$J_{4'',5''}$	1.1	4.32	$J_{4'',5''}$	1.5
H-5''	4.51			4.49		
acetyl	2.08			2.08		

Expression and purification of glycosyltransferase enzymes

The gene encoding *C. concisus* ATCC 33237 PglJ and *C. jejuni* PglJ were codon optimized for expression in *E. coli*, synthesized, and cloned into pET-29b (+) vector by Twist Bioscience (San Francisco, USA). The *C. concisus* strain 13826 PglJ gene (Twist Bioscience) was cloned into pET24a vector using Gibson Assembly. The proteins were expressed from *E. coli* BL21 (DE3) cells (Invitrogen) grown in autoinduction media supplemented with 150 µg/mL kanamycin. Cells were grown at 37 °C until OD_{600nm} of 0.8-1.0 was reached, followed by overnight incubation at 18 °C, while shaking at 220 rpm. Cells were harvested via centrifugation and resuspended in 5 mL of 50 mM HEPES pH 7.5, 200 mM NaCl, 5% glycerol, 1 mg DNase I, 5 mg lysozyme and 1 tablet of protease inhibitors (Pierce), per gram of cell pellet. The cells were lysed via microfluidization at 18,000 PSI and centrifuged at 106,000 x g for 30 minutes. The supernatant was discarded, and cell envelope fraction (CEF) was resuspended in 20 mL of membrane extraction buffer (50 mM HEPES, pH 7.5, 200 mM NaCl 5% glycerol, 1% Triton X-100, 25 mM imidazole). The pellet was homogenized using a Dounce tissue homogenizer and stirred at 4 °C for 1 hour. The membrane fraction was clarified via centrifugation at 106,000 x g for 30 minutes. The clarified lysate was run through a 1 mL Ni-NTA gravity column pre-equilibrated with 5 column volumes (CV) of equilibration buffer (50 mM HEPES, pH 7.5, 200 mM NaCl, 5% glycerol, 25 mM imidazole, 0.03% DDM). The column was then washed with 5 CV of the wash buffer (50 mM HEPES, pH 7.5, 200 mM NaCl, 5% glycerol, 50 mM imidazole, 0.03% DDM). The protein was eluted in two steps using 3 CV of elution buffer E1 (50 mM HEPES, pH 7.5, 200 mM NaCl, 5% glycerol, 100 mM imidazole, 0.03% DDM) and 3 CV of E2 (50 mM HEPES, pH 7.5, 200 mM NaCl, 5%

glycerol, 200 mM imidazole, 0.03% DDM). The highest purity fractions were pooled and desalted in 10 mM HEPES, pH 7.5, 150 mM NaCl, 5% glycerol, 2 mM DTT, 0.03% DDM using 3 x 5 mL Cytiva HiTrap columns. *C. jejuni* PglJ was desalted in 50 mM HEPES, pH 7.5, 150 mM NaCl, 5% glycerol, 2 mM DTT, 0.03% DDM. The proteins were concentrated using 10 kDa MWCO centrifugal concentrators and flash frozen in liquid nitrogen. The purify of proteins was analyzed using Coomassie-stained SDS-PAGE (**Figure S11**). Site-directed mutagenesis to produce the *C. concisus* 33237 PglJ$^{R122C/S120N}$ variant was performed using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs) and the following mutagenic primers: 5'- GTGCTCGTGCCCAAGTATCTTG-3'(forward) and 5' TCGTTGATAACAAGACGTTTCTTAAATC -3' (reverse).

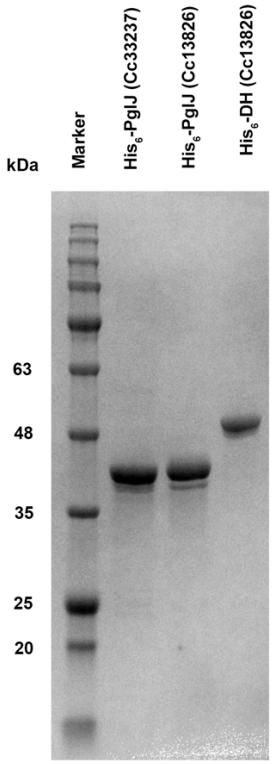


Figure S13. SDS-PAGE analysis of A) purified *C. concisus* 33237 PglJ, *C. concisus* 13826 PglJ, and *C. concisus* 13826 dehydrogenase (DH).

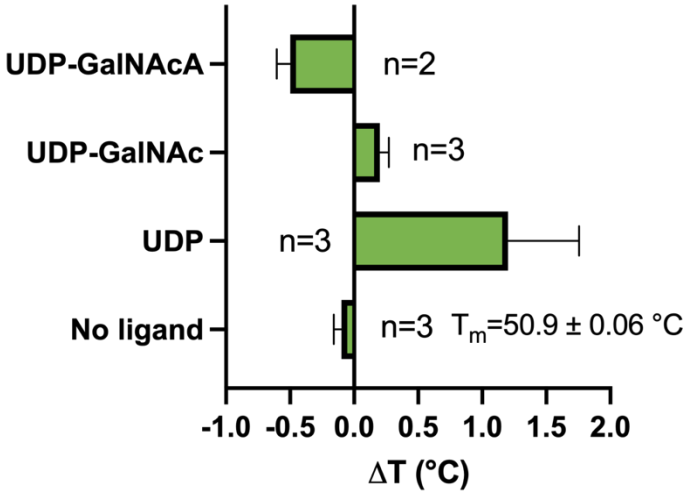


Figure S14. The nano differential scanning fluorimetry (NanoDSF) experiments are shown for S120N/R122C *Cc* PglJ 33237 tested against 1 mM UDP and a panel of UDP-sugars. Each data point represents the mean value derived from three independent measurements (n = 3) with error bars indicating standard deviation (± SD). The unliganded S120N/R122C *Cc* PglJ 33237 exhibited a melting point (Tm) of 50.9 ± 0.06 °C.

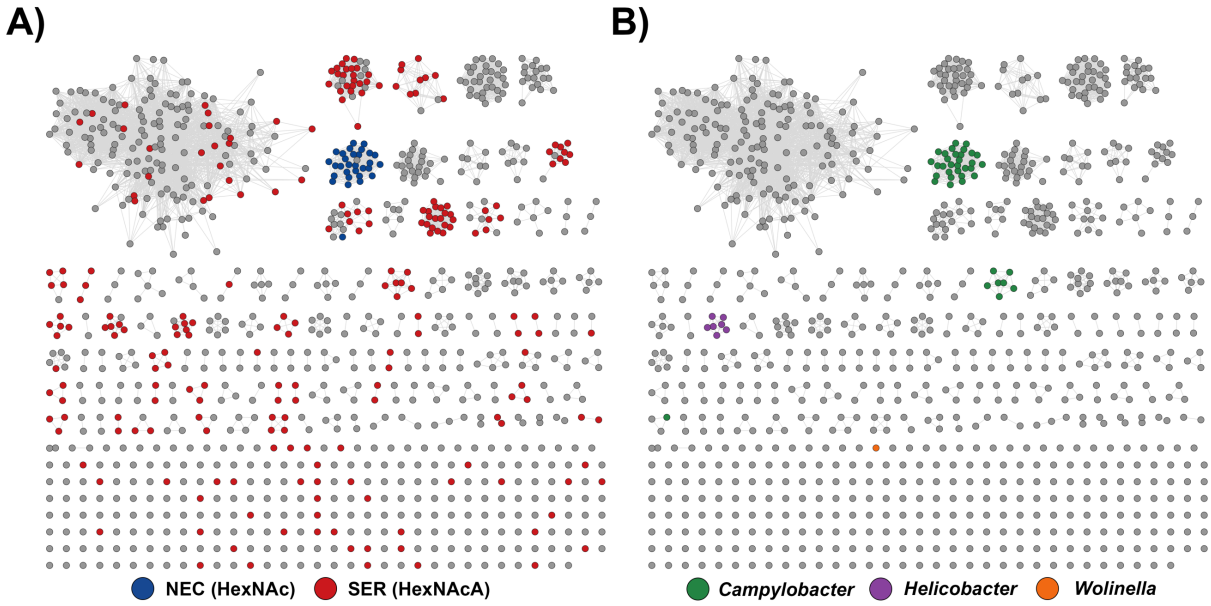


Figure S15. A) Visualized is the sequence similarity network (SSN) of PglJ homologs, where each node corresponds to a 90% representative sequence. Nodes containing sequences that bear NEC and SER motifs are highlighted with distinctive colors. B) The same SSN is recolored to showcase the distribution of PglJ orthologs within the genera *Campylobacter, Helicobacter, and Wolinella*. These proteins, inferred to maintain the same function as PglJ, were identified via their genomic positioning within the *pgl* operon (as revealed by Genome Neighborhood Diagrams (GNDs)) or by their high sequence identity with functionally characterized PglJs.

Table S3. *Campylobacter* PglJ Orthologs featuring the SER Motif and dehydrogenase orthologs implicated in UDP-GalNAcA synthesis. Certain PglJ orthologs identified in the SSN have been excluded due to their origin from incompletely annotated genomes, precluding the display of GNDs for the identification of the associated dehydrogenase in the operon. Out of an initial set of 37 sequences from annotated genomes, those with high redundancy (>95%) were excluded.

Organism	PglJ UniProt ID	Dehydrogenase UniProt ID
C. concisus (strain 13826)	A7ZET7	A7ZEU4
C. concisus (ATCC 33237)	A0A0M3V2L4	A0A0M5MEL7
Campylobacter pinnipediorum subsp. pinnipediorum	A0A1S6TPM8	A0A1T2XDH4
Campylobacter pinnipediorum subsp. caledonicus	A0A1S6U7H4	A0A1S6U7E4
Campylobacter concisus ATCC 51562	U2GB70	U2F4J9
Campylobacter curvus (strain 525.92)	A7GWW4	A7GWV7
Campylobacter massiliensis	A0A842JCJ2	A0A842JF46
Campylobacter showae RM3277/ATCC 51146	C6RIL1	C6RIL6
Campylobacter rectus RM3267/ATCC 33238	B9D371	B9D375

Table S4. The *Campylobacter* PglJ orthologs possessing both the NEC motif and having a co-localized dehydrogenase in the pgl operon. Out of an initial set of 6 sequences from annotated genomes, those with high redundancy (>95%) were removed.

Organism	PglJ UniProt ID	Dehydrogenase UniProt ID
Campylobacter iguaniorum	A0A076FB33	A0A076FCV9
Campylobacter fetus subsp. fetus	A0A0S4SCF2	A0A0S4SDK3
Campylobacter sputorum subsp. sputorum	A0A381DH92	A0A381DHA2
Campylobacter fetus	A0A5L4L5Y6	A0A5L8K203

Und-PP-diNAcBac-GalNAc enzymatic synthesis

The Und-PP-Bac-GalNAc reaction was set up in a 7 mL scintillation vial. The reaction contained a total volume of 1.5 mL and consisted of 267 μM UndP, 267 μM UDP-diNAcBac, 100 nM *Cc* PglC, 400 μM UDP-GalNAc, 100 nM *Cc* PglA, 50 mM HEPES pH 7.5, 100 mM NaCl, 0.1% Triton X-100 and 5 mM $MgCl_2$. The reactions contained a final concentration of 10% DMSO.

Buffers and solvents.

- PglC buffer: 50 mM HEPES pH 7.5, 100 mM NaCl, 0.03% DDM
- PglA buffer: 10 mM HEPES pH 7.5, 150 mM NaCl, 2 mM DTT, 0.03% DDM
- Assay buffer: 50 mM HEPES pH 7.5, 100 mM NaCl, 0.1% Triton X-100 and 5 mM $MgCl_2$
- Pure solvent upper phase (PSUP): 15 mL $CHCl_3$, 240 mL MeOH, 1.83 g KCl in 235 H_2O
- TLC solvent: 65:25:4 $CHCl_3$/MeOH/H_2O
- TLC stains: *CAM* – 0.5 g ceric ammonium sulfate, 12 g ammonium molybdate, 15 mL H_2SO_4, 235 mL H_2O
- Silica column: 0.3 g silica, 1 cm height, 1 cm diameter; pre-equilibrated in (7:1) $CHCl_3$/MeOH
- Column volume (CV): ~0.5 mL

Protocol.

First, UndP (150 μL of 2.67 mM in DMSO) and 1005 μL of assay buffer were combined and thoroughly mixed to distribute the DMSO. Then *Cc* PglC (30 μL of 5 μM) and *Cc* PglA (15 μL of 10 μM) was added to the reaction mixture. The reaction was initiated by the addition of UDP-diNAcBac (150 μL 2.67 μM in H_2O) followed by UDP-GalNAc (150 μL in 4 mM in H_2O) and allowed to proceed at ambient temperature for 30 min. After this time, the reaction was quenched with 2 mL of (2:1) $CHCl_3$/MeOH. Then 500 μL of PSUP was added and the solutions were vortexed and allowed to settle to form two layers, aqueous and organic. The upper, aqueous layer was removed and back-extracted with 1.5 mL of (2:1) $CHCl_3$/MeOH. The organic layers were combined. Then 500 μL of PSUP was added to the organic layer, vortexed, and allowed to settle. The aqueous layer was removed, and the organic layer was washed one more time with 500 μL PSUP. After removal of the

aqueous layer, the organic layer was concentrated under a stream of N_2, followed by a mini Na_2SO_4 pipette column to remove any remaining water. The eluted mixture was concentrated under N_2 to produce an oil. The oil was then re-suspended in a mixture of (7:1) $CHCl_3$/MeOH and loaded on a silica column. The mobile phase gradient is as follows: 3 CV (7:1) $CHCl_3$/MeOH, 5 CV (5:1) $CHCl_3$/MeOH, and 8 CV 100% MeOH. Each fraction (~0.5 mL) was analyzed by TLC and visualized with CAM staining. Subsequently, each fraction was quantified by the UDP-Glo biochemical assay.

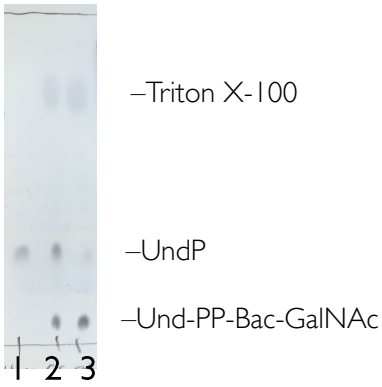


Figure S16. Thin-layer chromatography (TLC) of the enzymatic reaction to Und-PP-diNAcBac-GalNAc. The solvent system used was 64:25:4 $CHCl_3$/MeOH/H_2O and visualized with CAM staining. Lanes: 1) Und-P starting material, 2) Co-spot, and 3) reaction after extraction.

UDP-Glo biochemical assays to quantify Und-PP-diNAcBac-GalNAc

Und-PP-diNAcBac-GalNAc concentration determination assays were performed with *Cc* PglJ using the Promega UDP-Glo kit from Promega, which detects UDP generated over the course of the reaction. The quenching solution was prepared as described by Promega. A UDP-Glo standard curve was obtained using final [UDP] concentration of 50 µM, 25 µM, 12.5 µM, 6.25 µM, 3.125 µM, 1.56, and 0 µM from 10x UDP stocks in H_2O. The standard curve contained 10% DMSO. The PglJ assays contained 100 nM *Cc* PglJ, 0.1% Triton X-100, 50 mM HEPES at pH 7.5, 100 mM NaCl, 5 mM $MgCl_2$, 25 µM UDP-GalNAcA, and an unknown amount of Und-PP-diNAcBac-GalNAc in a final volume of 11 µL. An aliquot (5 µL) of each fraction from the Und-PP-diNAcBac-GalNAc purification was

placed in a 1.7 mL Eppendorf and concentrated using the SpeedVac Vacuum Concentrator (10 min). Und-PP-diNAcBac-GalNAc was resuspended in 1.1 µL of DMSO followed by the addition of assay buffer (7.7 µL). Then *Cc* PglJ (1.1 µL of 1 µM) was added to the reaction mixture lacking the UDP-sugar for 2 min at ambient temperature. The reactions were initiated by the addition of UDP-GalNAcA (1.1 µL of 250 µM in H_2O). The 11 µL reactions were quenched with 11 µL of the UDP detection reagent after 30 minutes. The reaction mixture (20 µL) was transferred to a 96-well plate (white, nonbinding surface, Corning). The plate was shaken at low speed for 30 s and incubated for 1 h at 25 ºC, and luminescence was read on the plate reader. All luminescence values were background subtracted before converting to UDP.

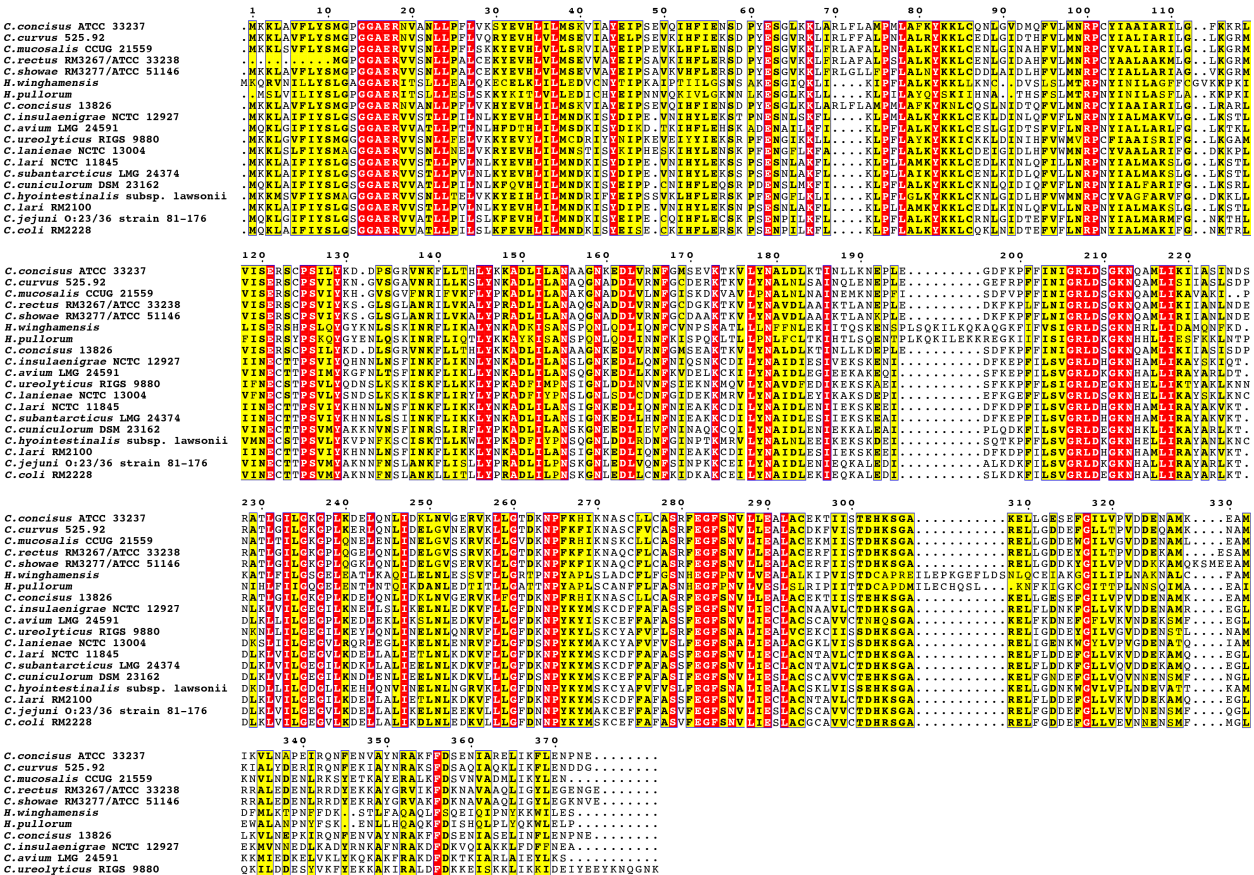


Figure S17. Full sequence alignment of PglJ glycosyltransferases from different Campylobacter species.

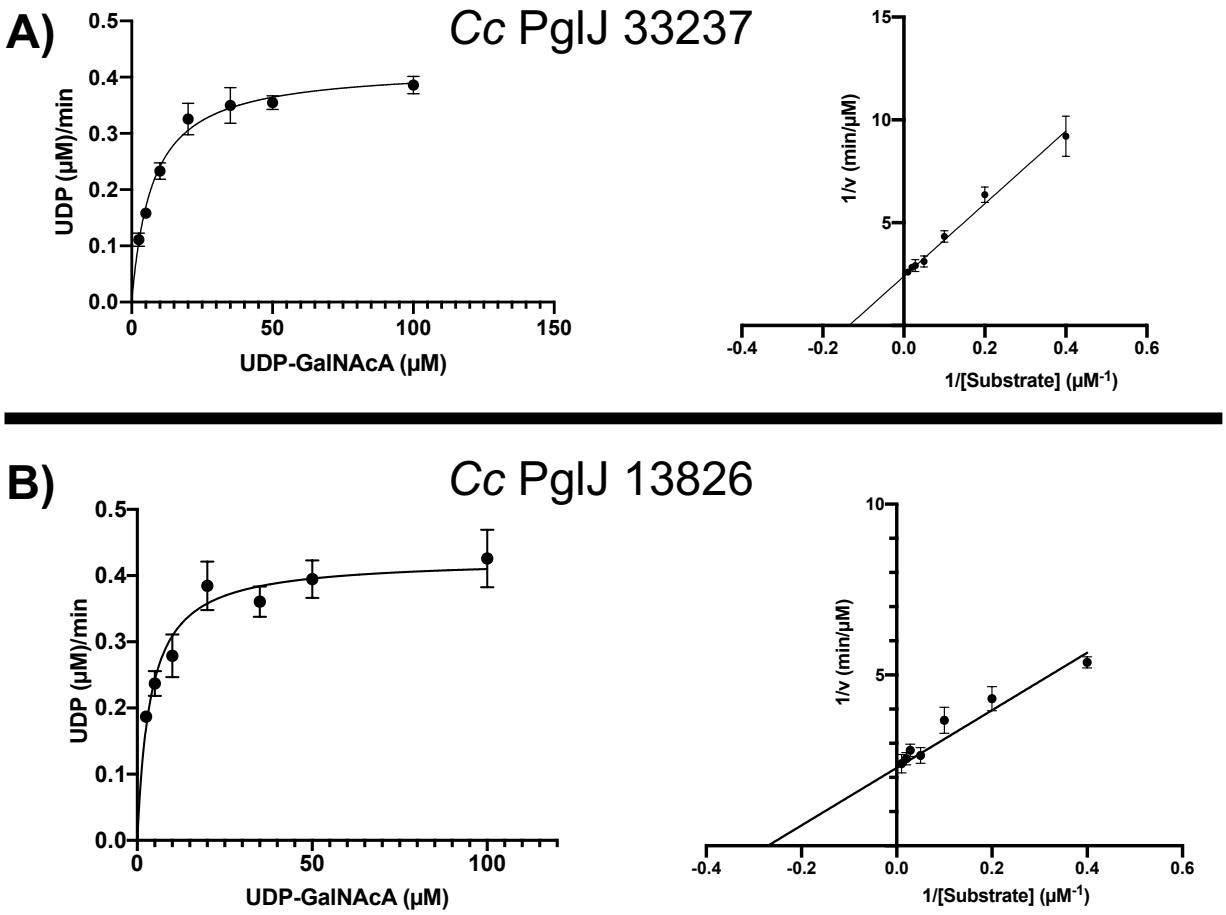


Figure S18. Kinetics of PglJ from two different strains of *C. concisus,* **A**) 33237 and **B**) 13826. Determination of V_{max}, K_m, and k_{cat} of UDP-GalNAcA which was derived from Michaelis-Menten non-linear regression and Lineweaver Burke plots. The steady-state velocity in the presence of 2.5, 5, 10, 20, 35, 50, and 100 μM UDP-GalNAcA. Error bars are given for mean ± SEM, n = 3.

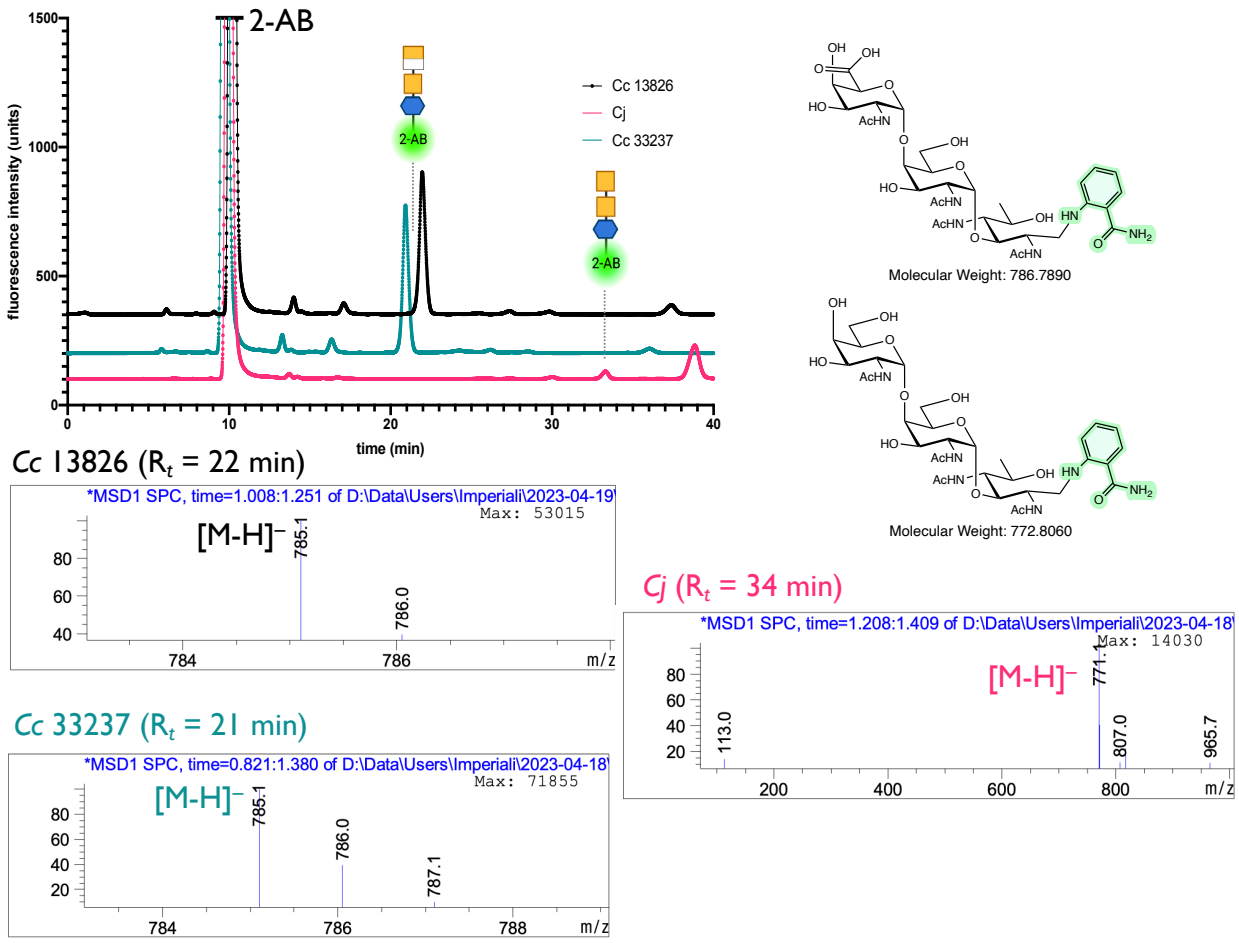
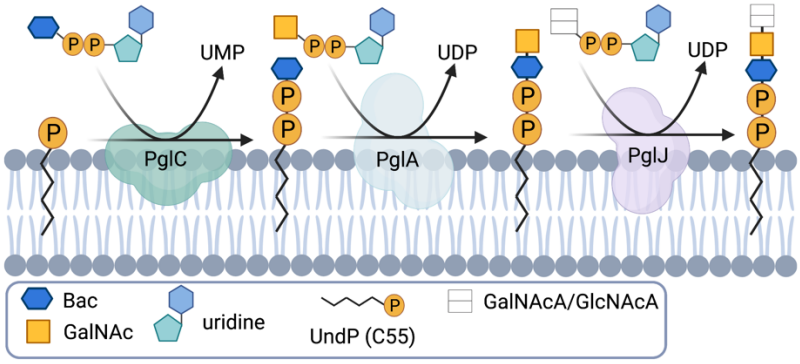


Figure S19. Characterization of the glycans produced from PglC, PglA, and PglJ from *C. concisus* (strains 33237 and 13826) and *C. jejuni* through 2-aminobenzamide labeling, fluorescence-based HPLC, and ESI-MS.



UDP Glo coupled-enzyme biochemical assays

C. concisus PglJ assays were performed using the Promega UDP-Glo assay, which detects UDP generated over the course of the reaction. The quenching solution was prepared as described by Promega. A UDP-Glo standard curve was obtained using final [UDP] concentrations of 10 µM, 5 µM, 2.5 µM, 1.25 µM, 0.625 µM, 0.3125 µM, 0.15625 µM, and 0 µM from 10x UDP stocks. The standard curve contained 10% DMSO. The PglJ assays contained 500 nM PglC, 0.5 nM PglA, 10 nM PglJ, 20 µM UndP (10% DMSO final), 0.1% Triton X-100, 50 mM HEPES at pH 7.5, 100 mM NaCl, 5 mM $MgCl_2$, 20 µM UDP-diNAcBac, 20 µM UDP-GalNAc, and 10 µM of UDP-GalNAcA in a final volume of 11 µL. PglC, PglA, and PglJ were preincubated in the reaction mixture lacking UDP-sugars for 5 min at ambient temperature. The reaction rate of PglA was predetermined to be linear over 8 min at the given concentrations. Upon the addition of the UDP-sugar mixture, the reaction was allowed to proceed for 4 min before the addition of quenching solution. PglA is kinetically sensitive under these reaction conditions. The reaction mixture was transferred to a 96-well plate (white, nonbinding surface, Corning). The plate was shaken at low speed for 30 s and incubated for 1 h at 25 ºC, and luminescence was read on the plate reader.

1	2	3	4	5	Conditions
PglC+PglA	PglC+PglA+PglJ	PglC+PglJ	PglC+PglA	PglC+PglA+PglJ	**Enzymes**
UDP-Bac + UDP-GalNAc		UDP-Bac + UDP-GalNAc + UDP-GalNAcA			**Sugars**

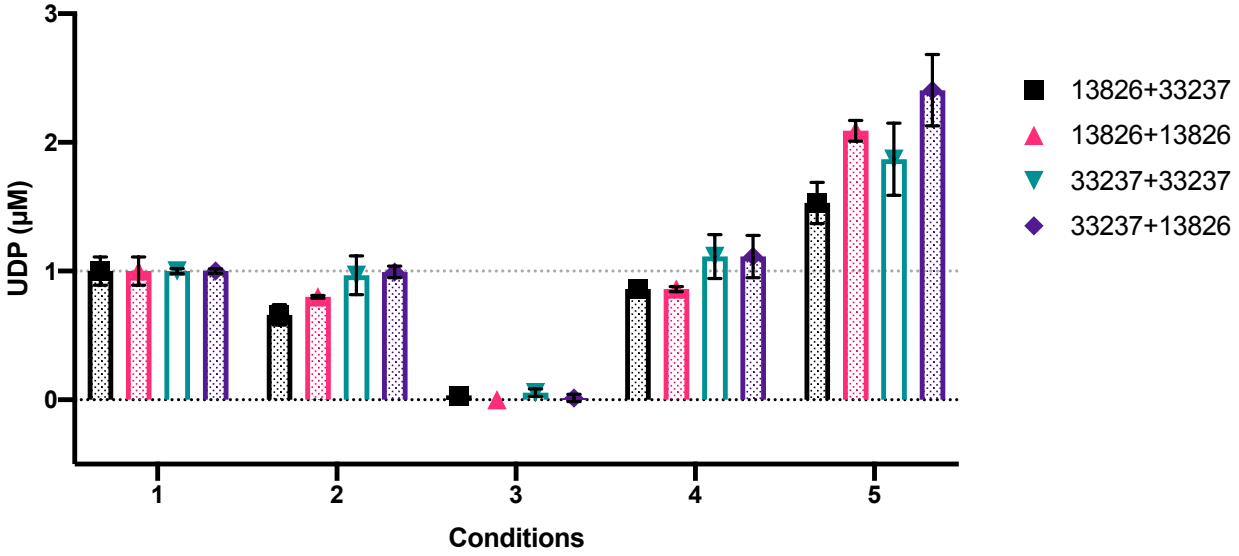


Figure S20. *C. concisus* PglC, PglA, and PglJ UDP-Glo coupled-enzyme assays. Error bars represent standard deviation of biological replicates (33237+33237 and 33237+13826, n=3; 13826+13826 and 13826+33237, n=4).

Expression and purification of *C. concisus* strain 13826 TviB family dehydrogenase (*Cc* DH)

The gene encoding WbpO was inserted into pMCSG7 vector using NEBuilder® HiFi DNA Assembly (**Table S3**). The N-terminally His_6-tagged protein was expressed in the same manner as glycosyltransferases (*vide supra*). After centrifugation, the cell pellet was resuspended in in 5 mL of 50 mM HEPES pH 7.5, 200 mM NaCl, 5% glycerol, 1 mg DNase I, 5 mg lysozyme and 1 tablet of protease inhibitors (Pierce), per gram of cell pellet. The cells were lysed via microfluidization at 18,000 PSI and centrifuged at 106,000 x g for 30 minutes. The clarified lysate was run through a 1 mL Ni-NTA gravity column pre-equilibrated with 5 column volumes (CV) of equilibration buffer (50 mM HEPES, pH 7.5, 200 mM NaCl, 5% glycerol, 25 mM imidazole). The column was then washed with 5

CV of the wash buffer (50 mM HEPES, pH 7.5, 200 mM NaCl, 5% glycerol, 50 mM imidazole). The protein was eluted in two steps using 3 CV of elution buffer E1 (50 mM HEPES, pH 7.5, 200 mM NaCl, 5% glycerol, 100 mM imidazole) and 3 CV of E2 (50 mM HEPES, pH 7.5, 200 mM NaCl, 5% glycerol, 200 mM imidazole). The protein was desalted in 10 mM HEPES, pH 7.5, 150 mM NaCl, 5% glycerol, 2 mM DTT, using 3 x 5 mL Cytiva HiTrap columns.

Table S5. Primers used for generating overlapping insert and vector fragments in Gibson Assembly of *C. concisus* 13826 TviB family dehydrogenase (*Cc*DH) and pMCSG7 vector.

Primer	Sequence (5'-3')
*Cc*DH Forward	AATCCAATATTGGAAGTGGAATGAAGATTGCCGTGGTTG
*Cc*DH Reverse	GCTTTGTTAGCAGCCGGATCTTATAAACGCGCATCTGC
pMCSG7 Forward	GATCCGGCTGCTAACAAAG
pMCSG7 Reverse	TCCACTTCCAATATTGGATTG

Campylobacter concisus dehydrogenase (13826) assays

Assays that contain magnesium chloride ($MgCl_2$): An aliquot of NAD^+ (6.78 µL of 36.9 mM in H_2O) was placed in a 0.65 mL micro centrifuge tube and thoroughly mixed with assay buffer (71.72 µL of 14 mM Tris, pH 8.2, 6.97 mM $MgCl_2$). The *C. concisus* dehydrogenase (9 µL of 0.46 mM) was added and allowed to incubate at ambient temperature for 5 min. Then UDP-GalNAc or UDP-GlcNAc (10 µL of 10 mM) was added to the reaction mixture and incubated at 37 ºC. The *Cc* DH assays contained final concentrations of 41.4 µM *Cc* DH, 10 mM Tris at pH 8.2, 5 mM $MgCl_2$, 1 mM UDP-sugar, and 2.5 mM NAD^+ in a final volume of 100 µL.

Assays that contain ammonium sulfate ($(NH_4)_2SO_4$): An aliquot of NAD^+ (6.78 µL of 36.9 mM in H_2O) was placed in a 0.65 mL micro centrifuge tube and thoroughly mixed with assay buffer (74.22 µL of 0.135 mM Tris, pH 8.2, 0.135 mM $(NH_4)_2SO_4$). The *C. concisus* dehydrogenase (9 µL of 0.46 mM) was added and allowed to incubate at ambient temperature for 5 min. Then UDP-GalNAc or UDP-GlcNAc (10 µL of 10 mM) was added to the reaction mixture and incubated at 37 ºC. The *Cc* DH assays contained final concentrations of 41.4 µM *Cc* DH, 100 mM Tris at pH 8.2, 100 mM $(NH_4)_2SO_4$, 1 mM UDP-sugar, and 2.5 mM NAD^+ in a final volume of 100 µL.

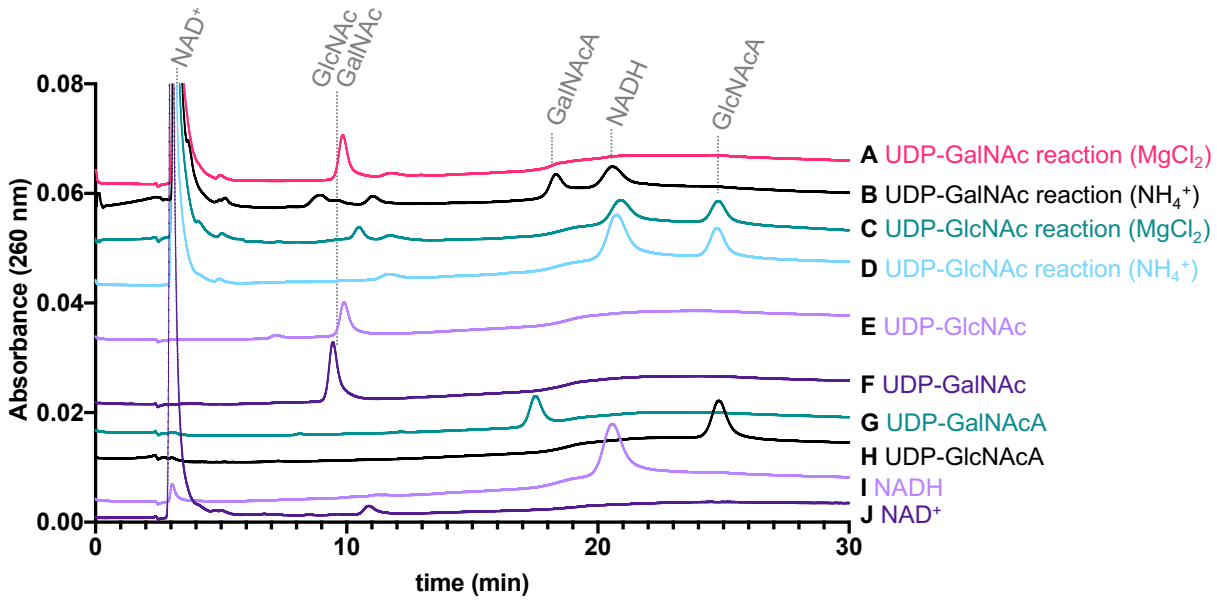


Figure S21. Monitoring the *C. concisus* dehydrogenase reactions by HPLC at 260 nm. A) UDP-GlcNAc reaction, B) UDP-GalNAc reaction, C) UDP-GalNAcA standard, D) UDP-GlcNAc standard, E) UDP-GlcAcA standard, F) UDP-GalNAc standard, G) NAD^+ standard, and G) NADH standard. HPLC-based assay monitoring the production of oxidized UDP-sugars in the presence of $(NH_4)_2SO_4$ and $MgCl_2$. Reaction lanes were Y-axis-shifted using GraphPad Prism 8 to allow for better visualization of reaction traces with a relative absorbance at 260 nm.

References

(1) Varki, A.; Cummings, R. D.; Aebi, M.; Packer, N. H.; Seeberger, P. H.; Esko, J. D.; Stanley, P.; Hart, G.; Darvill, A.; Kinoshita, T.; et al. Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **2015**, *25* (12), 1323-1324. DOI: 10.1093/glycob/cwv091.

(2) Rejzek, M.; Mukhopadhyay, B.; Wenzel, C. Q.; Lam, J. S.; Field, R. A. Direct oxidation of sugar nucleotides to the corresponding uronic acids: TEMPO and platinum-based procedures. *Carbohydr. Res.* **2007**, *342* (3), 460-466. DOI: https://doi.org/10.1016/j.carres.2006.10.016. Larkin, A.; Imperiali, B. Biosynthesis of UDP-GlcNAc(3NAc)A by WbpB, WbpE, and WbpD: Enzymes in the Wbp Pathway Responsible for O-Antigen Assembly in *Pseudomonas aeruginosa* PAO1. *Biochemistry* **2009**, *48* (23), 5446-5455. DOI: 10.1021/bi900186u.

(3) Miller, W. L.; Matewish, M. J.; McNally, D. J.; Ishiyama, N.; Anderson, E. M.; Brewer, D.; Brisson, J.-R.; Berghuis, A. M.; Lam, J. S. Flagellin Glycosylation in *Pseudomonas aeruginosa* PAK Requires the O-antigen Biosynthesis Enzyme WbpO. *J. Biol. Chem.* **2008**, *283* (6), 3507-3518. DOI: https://doi.org/10.1074/jbc.M708894200.