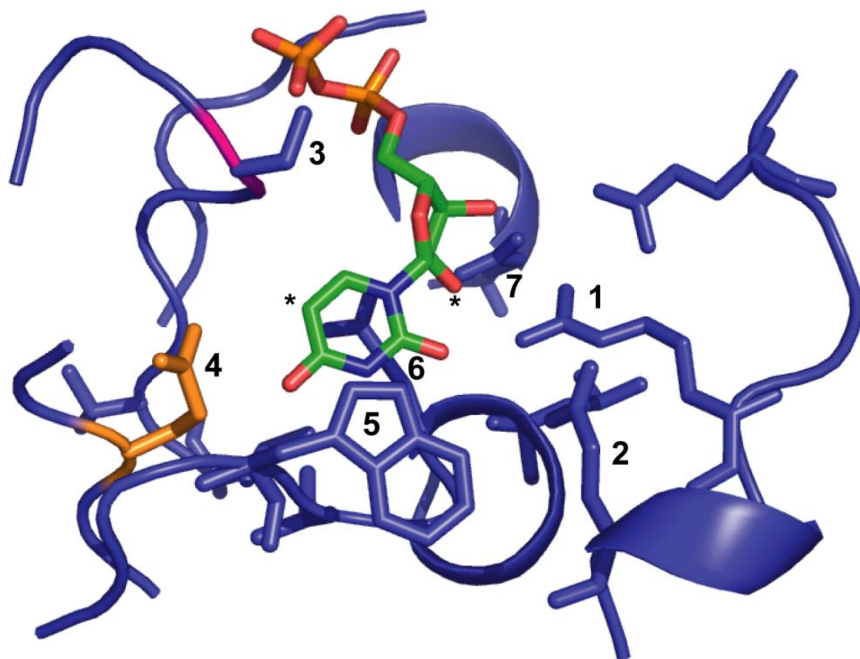
An *Escherichia coli* strain for whole-cell based glycodiversification of natural products (supplementary information)

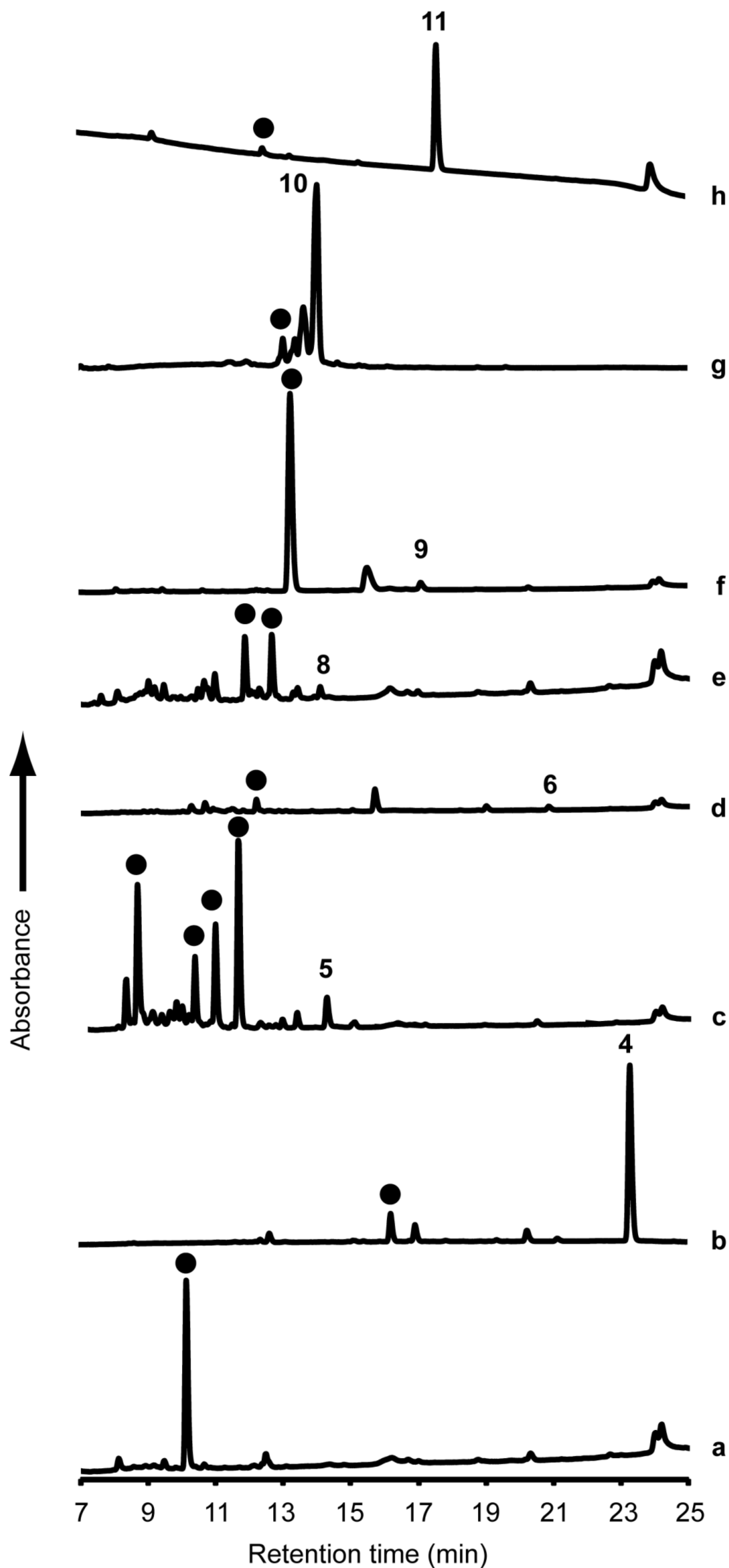
Gavin J Williams, Jie Yang, Changsheng Zhang , & Jon S Thorson

Laboratory for Biosynthetic Chemistry, Pharmaceutical Sciences Division, School of Pharmacy, National Cooperative Drug Discovery Program, University of Wisconsin-Madison, 777 Highland Avenue, Madison.

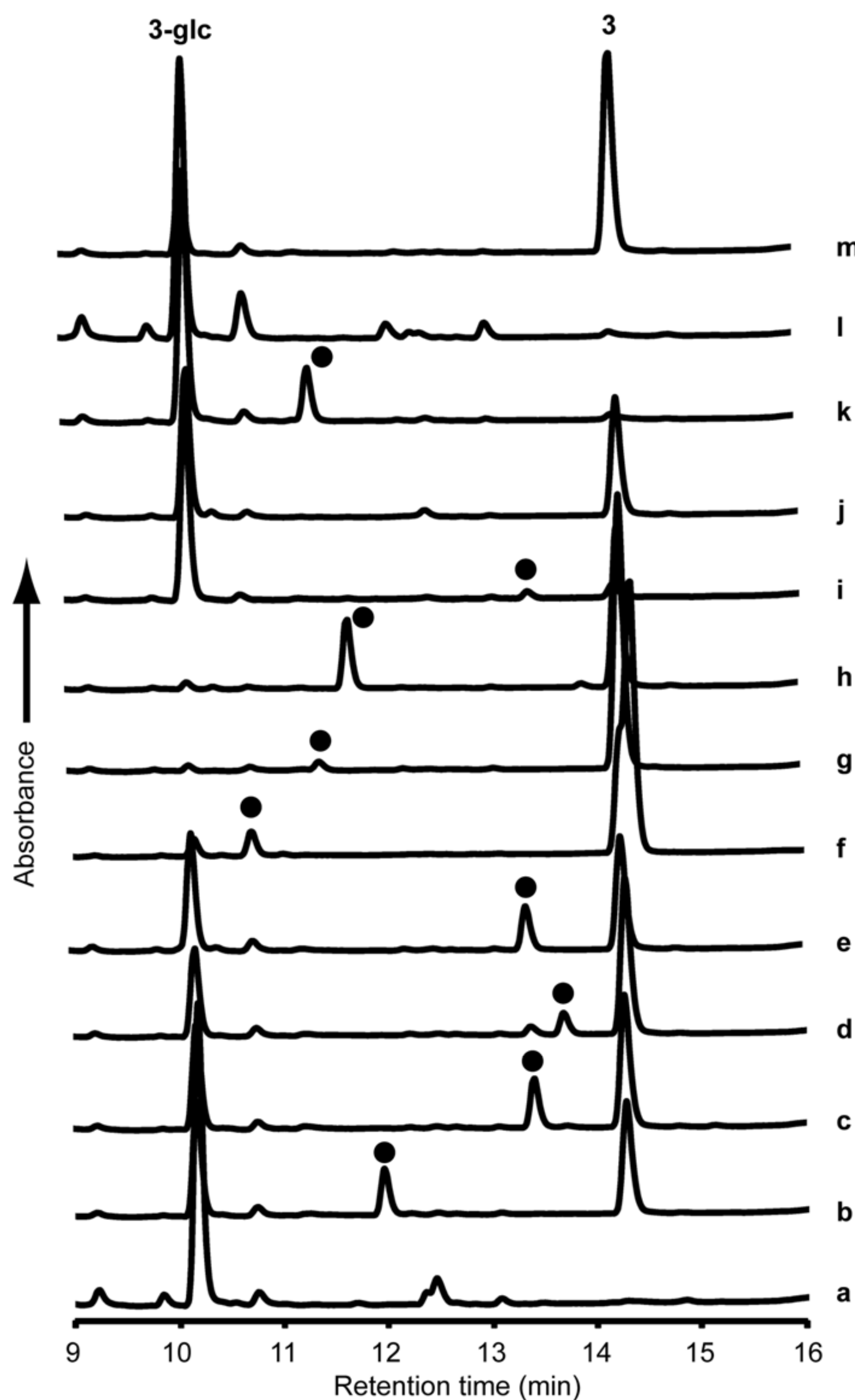
Supplementary Figures



Supplementary Figure 1. Location of residues selected for saturation mutagenesis in the UDP binding site of OleD (PDB ID 21YF). Each site for mutagenesis is shown numbered: 1, R219; 2, E222; 3, S237; 4, Q268; 5, W290, 6, Q293; 7, E314. Position 268 is also highlighted orange for clarity. UDP is illustrated as a stick model, asterisks indicate atomic position that differ in structure between UDP and TDP.



Supplementary Figure 2. OleD TDP16 bioconversion of acceptors **3-11**. (a) acceptor **3**; (b) acceptor **4**; (c) acceptor **5**; (d) acceptor **6**; (e) acceptor **8**; (f) acceptor **9**; (g) acceptor **10**; (h) acceptor **11**. It should be noted that with certain acceptors, OleD catalyzes iterative/multiple glycosylation.



Supplementary Figure 3. OleD 'TDP16' bioconversion of acceptor **3** with various free sugars. (a) D-glc, **12**; (b) 6-deoxy-glc, **13**; (c) 6-azido-glc, **14**; (d) 6-bromo-glc, **15**; (e) 6-chloro-glc, **16**; (f) 4-deoxy-glc, **17**; (g) 3-fluoro-glc, **20** (h) 2-deoxy-glc, **18**; (i) 2,6-dideoxy-glc, **19**; (j) D-gal, **22**; (k) D-xyl, **23**; (l) N-acetylglucosamine, **21**; (m) no sugar. Each successful glycoside is identified by a filled circle.

Supplementary Table 1. Amino acid mutations of enzyme variants used in this study.

Enzyme variant	Amino acid mutations	Reference
GalK	M173L/Y371H	Yang et al., Chem Biol 2005 [1]
RmlA L89T	L89T	Barton et al, PNAS 2002 [2]
OleD ‘ASP’	P67T/S132F/A242V	Williams et al., Nat Chem Biol 2007 [3]
OleD 3-1H12	P67T/S132F/A242L	Williams et al., Nat Prot 2008 [4]
OleD-1C9	P67T/I112K/A242V	Williams et al., Chem Biol 2008 [5]
OleD-TDP16	P67T/S132F/A242L/Q268V	This study

Supplementary Table 2. Kinetic parameters of WT and mutant OleD toward **3** as acceptor and **1** or **2** as donor.

Enzyme variant	Donor	k_{cat} (min^{-1})	K_m (mM)	k_{cat}/K_m ($mM^{-1}.min^{-1}$)	k_{cat}/K_m(**2**)/ k_{cat}/K_m(**1**)
WT	**1**	-[a,b]	-[a,b]	0.18	NA
WT	**2**	-[a,b]	-[a,b]	0.04	0.22
ASP	**1**	2.6 ±0.04	0.07 ±0.006	37	NA
ASP	**2**	-[a,b]	-[a,b]	6.3	0.17
3-1H12	**1**	17.2 ±0.5	0.27 ±0.03	64	NA
3-1H12	**2**	14.6 ±2.0[b]	0.63 ±0.17[b]	47	0.73
TDP16	**1**	47.1 ±1.9	0.84 ±0.14	56	NA
TDP16	**2**	29.3 ±1.7	0.44 ±0.1	67	1.20

[a]could not be saturated with donor, k_{cat}/K_m estimated from linear fit of data

[b]strong substrate inhibition, data below 1.5 mM **2** used to fit to Michealis-Menton curve

Supplementary Table 3. Donor promiscuity of TDP16.

Donor	Product retention time (mins)[a]	Conversion rate (%)[b]	
		ASP[c]	TDP16
UDP-glc	10.04	96	60
UDP-2-deoxy-glucose	N.D	7	N.D
UDP-3-deoxy-glucose	10.74	8	20
UDP-4-deoxy-glucose	10.58	22	26
UDP-4,6-dideoxy-glucose	13.16	45	33
UDP-6-deoxy-glucose	11.58	85	31
UDP-xylose	11.24	31	41
UDP-3-azido-glucose	ND	ND	ND
UDP-4-azido-glucose	ND	ND	ND
UDP-6-azido-glucose	13.37	41	41
UDP-glucosamine	9.31	24	28
UDP-3-amino-glucose	9.82	4	7.4
UDP-4-amino-glucose	10.08	13	21
UDP-6-amino-glucose	9.04	3	17
UDP-3-amino-6-deoxy-glucose	11.08	4	3
UDP-4-amino-6-deoxy-glucose	10.78	47	39
UDP-galactose	9.79	ND	1.4
UDP-mannose	ND	ND	ND
UDP-talose	ND	ND	ND
UDP-3-acetamido-glucose	ND	ND	ND
UDP-3*O*-methyl-glucose	ND	ND	ND
UDP-2-acetamido-glucose	9.24	21	31
UDP-6-thio-glucose	13.00	47	20

[a]See Experimental Procedures for details of HPLC conditions.

[b]Percent conversions were determined by HPLC and calculated by dividing the integrated area of the glycosylated product by the sum of the integrated area of the product plus the integrated area of the remaining acceptor **4**. Product retentions were consistent with previous values. See Supplementary Methods for reaction conditions.

[c]Data from ref [3]

ND, non-detected (estimated minimal detection limit of 0.1% conversion from **3**).

Supplementary Table 4. ASP and TDP16 in vitro glucosylation rates with panel of acceptors.

Acceptor/ donor[a]	Product retention time (mins)[b]	Conversion[c]			MS (m/z)	
		ASP	TDP16	Fold improvement[d]	calcd	found
3/2	10.1	19	87	4.6	338.1	$[M+H]^+$ 339.2
4/2	19.7	3	39	13	557.2	$[M+H]^+$ 558.2
5/2		23	82	3.6	416.1	$[M+H]^+$ 417.0
7/2	20.6	11	44	4	606.3	$[M+H]^+$ 607.2
10/2	21.6	1.5	4.8	3.2	1087.6	$[M-H]^-$ 1086.4
11/2	16.4	6.5	13	2	536.3	$[M+Cl]^-$ 571.2

[a]See Figure 2 for structures.

[b]See Experimental Procedures for details of HPLC conditions. Refers to major product.

[c]Percent conversions were determined by HPLC and calculated by dividing the integrated area of the glycosylated product by the sum of the integrated area of the product plus the integrated area of the remaining acceptor. See Experimental Procedures for reaction conditions.

[d]Fold improvement of TDP16 compared to ASP.

Supplementary Table 5. Percent conversion of acceptors **3-11** to glucosides.

	Strain/treatment[b]						MS (m/z)	
Acceptor[a]	WT/ no treatment	TDP16/ no treatment	1C9/ no treatment	TDP16/ detergent	TDP16/ sonication	ΔIlp/ no treatment	calcd	found
3	64	100	-	100	100	100	338.1	$[M+H]^+$ 339.2
4	1.2	12.2	84	36	22	32	557.2	$[M+H]^+$ 558.2
5	52	100	-	100	100	100	416.1	$[M+H]^+$ 417.2
6	51	55	-	11	23	35	350.1	$[M-H]^-$ 348.6; $[M+Cl]^-$ 384.8
7	ND	ND	-	ND	-	ND	606.3	ND
8	62	100	-	100	100	100	390.1	$[M+Cl]^-$ 425.0
9	6.4	97	-	99	100	99	323.0	$[M+H]^+$ 324.0
10	25	23	-	15.6	23.2	24	1087.6	$[M-H]^-$ 1086.4
11	2.5	4.9	-	5.6	5.1	3.8	536.3	ND

[a]See Figure 2 for structures.

[b]See Experimental Procedures for details of strains, treatments and HPLC conditions

Supplementary Table 6. In vivo glucosylation of **3** and sugars **12-23**.

Sugar[a]	Strain/treatment[b]						MS (m/z)	
	WT/ no treatment	TDP16/ no treatment	w/o GalK/RmlA no treatment	TDP16/ detergent	TDP16/ sonication	Δllp/ no treatment	calcd	found
12	91.0	99.0	95.6	95.0	98.4	98.0	338.1	$[M+H]^+$ 339.2
13	2.5	15.4	ND	11.5	11.7	5.5	322.1	$[M+H]^+$ 323.2
14	ND	16.6	ND	18.2	17.8	8.9	363.11	$[M+H]^+$ 364.2
15	1.0	7.5	ND	8.5	7.6	3.4	400.0	$[M+H]^+$ 401.0
16	3.3	16.1	ND	9.7	13.8	6.3	356.1	$[M+H]^+$ 357.0
17	ND	5.1	ND	ND	ND	ND	322.1	$[M+Cl]^-$ 357.0
18	ND	27.4	ND	22.0	21.2	29.4	322.1	$[M+Cl]^-$ 357.0
19	ND	3.2	ND	ND	ND	ND	306.11	$[M+Cl]^-$ 341.0
20	ND	2.5	11.9	3.1	2.1	2.5	340.1	$[M+H]^+$ 341.2
21	ND	ND	ND	ND	ND	ND	378.1	N.D
22	ND	ND	ND	ND	ND	ND	338.1	N.D
23	ND	17.2	3.5	14.6	9.5	4.2	308.1	309.2

[a]See Figure 2 for structures.

[b]See Experimental Procedures for details of strains, treatments and HPLC conditions.

Supplementary Methods

Construction of pDuet-GalK-Ep. The nucleotidylyltransferase (Ep) gene rmlA was ligated into *Eco*RI/*Nde*I digested pET-28a to give the plasmid pET28a-Ep, designed to express N-His$_6$-Ep. The QuikChange II Site-Directed Mutagenesis Kit (Stratagene) was used for mutagenesis of the target plasmid pET28a-Ep. The *Nde*I restriction site was eliminated by using QuikChange II Site-Directed Mutagenesis Kit (Stratagene) from pET28a-Ep template, using a pair of suitable mutagenic primers (5'-GCCGCGCGGCAGCAATATGAAAACGCGTAAGG-3' and 5'-CCTTACGCGTTTTCATATTGCTGCCGCGCGGC-3'). The desired point mutation after *Dpn*I treatment and transformation was verified by sequencing to give plasmid pEpdNDEI. The plasmid pEpdNDEI was subsequently treated with *Xba*I and *Eco*RI (Promega) and ligated into the vector pETDuet-1 (Novagen) to produce plasmid pETDuet-Ep. *E. coli* BL21 (DE3) cells were transformed with pETDuet-Ep. In a similar fashion, plasmid pGalKKpnI (which lacks a stop codon and contains a new *Kpn*I site) was constructed from the pGalKMLYH template using a pair of suitable mutagenic primers (5'-GCAGGACAGTGCTGCGGTACCGGCTGCTAACAAAGC-3' and 5'-GCTTTGTTAGCAGCCGGTACCGCAGCACTGTCCTGC-3'). The desired point mutation after *Dpn*I treatment and transformation was verified by sequencing to give the plasmid pGalKKpnI which was subsequently treated with *Nde*I (Promega) and *Kpn*I (New England BioLabs) and ligated into the vector pETDuet-Ep to produce plasmid pETDuet-GalK-Ep.

Construction of pCDF-TDP16. Plasmid pET28a-TDP16 was digested with *Nde*I and *Hin*dIII and the resulting fragment was ligated into similarly digested pCDF (Novagen) to afford pCDF-TDP16.

Saturation library preparation. The saturation mutagenesis libraries 'P67X', 'I112X', and 'A242X' were constructed using the Stratagene QuikChange II Site-Directed Mutagenesis Kit, as described by the manufacturer using the mutant P67T/I112T/A242V as template. Each library plasmid DNA (which had been digested with *Dpn*I) was first transformed into Novablue chemical competent cells. Colonies from each library were pooled from the plates and used to inoculate 5 ml of LB medium supplemented with 50 μg/ml kanamycin for overnight growth. Then, plasmid was prepared from each culture and was used to transform into *E.coli* BL21(DE3)pLysS.

Screening. Individual colonies were used to inoculate wells of a 96-deep well microtitre plate wherein each well contained 1 ml of LB medium supplemented with 50 μg/ml kanamycin. Culture plates were tightly sealed with AeraSeal™ breathable film (Research Products International Corp.). After cell growth at 37°C for 18 h with shaking at 350 rpm, 100 μl of each culture was transferred to a fresh deep-well plate containing 1 ml of LB medium supplemented with 50 μg/ml kanamycin. The original plate was sealed and stored at 4°C, or a glycerol copy made by mixing 100 μl of each culture with 100 μl 50% (v/v) glycerol and storing at -80°C. The freshly inoculated plate was incubated at 37°C for 2-3 h with shaking at 350 rpm. Protein expression was induced at OD_{600} ~0.7, and isopropyl β-D-thiogalactoside (IPTG) was added to a final concentration of 0.4 mM and the plate incubated for 18 h at 18°C. Cells were harvested by centrifugation at 3000 *g* for 10 min at 4°C, the cell pellets thoroughly resuspended in chilled 50 mM Tris-HCl (pH 8.0) containing 10 mg/ml lysozyme (Sigma), and the plates were subjected to a single freeze/thaw cycle to lyse the cells. Following thawing, cell debris was collected by centrifugation at 3000 *g* for 20 min at 4°C and 50 μl of the cleared supernatant used for enzyme assay.

For the assay, cleared supernatant was mixed with an equal volume (50 μl) of 50 mM Tris-HCl (pH 8.0) containing 10 mM $MgCl_2$, 200 μM **3**, and 1.0 mM **2** using a Biomek FX Liquid Handling Workstation (Beckman Coulter, Fullerton, CA). Upon mixing, the reactions were incubated for 3 h at 30°C, at which point, the crude reaction mixture was mixed with an equal volume of MeOH, and centrifuged at 3000 *g* for 20 min at 4°C. Aliquots of each product mixture (40 μl) were analyzed by RP-HPLC as described below for determination of specific activity.

Protein expression and purification. For characterization of specific OleD variants, single colonies were used to inoculate 3 ml LB medium supplemented with 50 μg/ml kanamycin and cultured overnight at 37°C. The entire starter culture was then transferred to 1 liter LB medium supplemented with 50 μg/ml kanamycin and grown at 37°C until the OD_{600} was ~0.7, then IPTG to a final concentration of 0.4 mM was added and the flask incubated for 18 h at 18°C. Cell pellets were collected by centrifugation at 10,000 *g* and 4°C for 20 min, resuspended into 10 ml 20 mM phosphate buffer, pH 7.4, containing 0.5M NaCl and 10 mM imidazole and were lysed by sonication. Cell debris was removed by centrifugation at 10,000 *g* and 4°C for 30 min and the cleared supernatant immediately applied to 2 ml of nickel-nitrilotriacetic acid (Ni-NTA) resin (QIAgen Valencia, CA), pre-equilibrated with the lysis buffer. Protein was allowed to bind for 30 min at 4°C with gentle agitation, and the resin washed 4 times with 50 ml each lysis buffer. Finally, the enzyme was eluted by incubation of the resin with 2 ml lysis buffer containing 100 mM imidazole for 10 min at 4°C with gentle agitation. The purified enzyme was applied to a PD-10 desalting column (Amersham Biosciences AB) equilibrated with 50 mM Tris-HCl (pH 8.0) and eluted as described by the manufacturer. Protein aliquots were immediately flash frozen in liquid nitrogen and stored at -80°C. Protein purity was verified by SDS-PAGE. Protein quantification was carried out using the Bradford Protein Assay Kit from Bio-Rad.

Determination of kinetic parameters. Enzyme assays were carried out in a total volume of 100 μl 50 mM Tris-HCl (pH8.0) containing 5 mM $MgCl_2$, and typically 10-50 μg pure enzyme. Kinetic parameters k_{cat} and K_m were determined with either **1** and **2** as variable substrates. For the determination of K_m for **1**, **3** was constant at 5 mM and **1** was varied between 0.025 and 5 mM. For the determination of K_m for **2**, **3** was constant at 5 mM and **2** was varied between 0.05 and 25 mM. Each experiment was performed in triplicate. Aliquots (100μl) were removed between 0 and 30 min, at which time product formation was still linear with respect to time, and quenched with 100 μl of ice-cold MeOH, and centrifuged at 10,000 *g* for 10 min. Supernatants were analyzed by analytical reverse-phase HPLC with a Gemini 5μ C18 column (Phenomenex) using a gradient of 10-90% CH_3CN in 0.1% trifluoroacetic acid (TFA)/H_2O in 20 min, with detection at 254 nm. HPLC peak areas were integrated, and the product concentration calculated as a percent of the total peak area. The substrate **3** and the glucoside product **3-glc** HPLC peak areas were integrated, and the product concentration calculated as a percent of the total peak area. Initial velocities were fitted to the Michaelis-Menten equation using Sigma Plot.

Donor specificity. The total volume was 50 μl and the acceptor **3** was at 50 μM and the NDP-sugar at ~500 μM. The NDP-sugars were used directly from RmlA-catalyzed reactions [2, 6, 7]. Reactions were incubated at 25°C for 3 h. Aliquots (25 μl) were quenched with 25 μl of ice-cold MeOH, and centrifuged at 10,000 *g* for 10 min. Supernatants were analyzed by analytical reverse-phase HPLC as described above. HPLC peak areas were integrated, and the product concentration calculated as a percent of the total peak area. All products were characterized by LC-MS.

In vivo bioconversions. For in vivo glucosylation of acceptors, a starter culture of BL21(DE3) pCDF-TDP16 or other control strain was used to inoculate a suitable volume of LB media containing 50 μg/ml streptomcyin and grown at 37 °C with shaking. Expression was induced by the addition of 0.1 mM IPTG when the OD600 was ~0.6, and the cells were then incubated at 18 °C with shaking for 18 hrs. Cells were then washed four times with 10 x volume phosphate buffered saline (PBS) at 4 °C. Finally, cells were resuspended in a volume of PBS such that the OD_{600} was 7.0. Acceptor stock solutions (in DMSO) were added to suitable volume of cells to give 100 μM each of **3**-**9**, 1 mM **10**, and 0.2 mM **11** and the cell suspensions continued to incubate at 18 °C with rotation. Aliquots (100 μl) were removed at timely intervals. Cells were collected by centrifugation and the resulting supernatants analyzed directly by HPLC as described in the Supplementary Materials and Methods.

For in vivo glycosylation of **3** with non-natural sugars, a starter culture of BL21(DE3) pDuet-GalK-Ep pCDF-TDP16 or other control strain was used to inoculate a suitable volume of LB media containing 50 μg/ml ampicillin and 50 μg/ml streptomcyin and then grown at 37 °C with shaking. Expression was induced by the addition of 0.1 mM IPTG when the OD_{600} was ~0.6, and the cells were then incubated at 18 °C with shaking for 18 hrs. Cells were then washed four times with 10 x volume phosphate buffered saline at 4 °C. Finally, cells were resuspended in a volume of PBS such that the OD_{600} was 7.0. Acceptor **3** (in DMSO) was added to suitable volumes of cell suspension, and 100 mM stock solutions of each sugar **12-23** added to a final concentration of 4 mM. Aliquots (100 μl) were removed at timely intervals. Cells were collected by centrifugation and the resulting supernatants analyzed directly by HPLC as described above.

Physical disruption. In order to test the effects of physical disruption, after the final resuspension of cells, the sample was sonicated (2 x 30 sec pulses) on ice and used without clarification. In order to test the effect of detergent on permeabilization, cells after the final resuspension were also treated with Triton X-100 (0.1% v/v).

Gene deletion. Deletion of the llp gene was performed using the Quick & easy E.coli Gene Deletion Kit from Gene Bridges (Heidelberg, Germany) as described by the manufacturer's instructions. The oligonucleotide 'llp for' (5'-CTTTGTGTAATACTTGTAACGCTACATGGAGATTAACTCAATCTAGAGGGAATTAACCCTCACTAAAGGGCG-3') and 'llp rev' (5'-CGTGACGCAGTAGCGGTAAACGGCAGACAAAAAAAATGGCGCACAATGTGTAATACGACTCACTATAGGGCTC-3') were used to replace the *llp* gene of BL21(DE3) with a FRT-flanked kanamycin resistance cassette, which was then removed by recombination. DNA sequencing of successful recombinants was performed using the oligonucleotides 'llp FOR seq' (5' -GAATCC GATGGAAGCATCCTGT-3') and 'llp REV seq' (5'-CTTCAGTAGAGTCAGCGCAGAATG-3').

Supplementary References

1. Yang, J., Fu, X., Liao, J., Liu, L. & Thorson, J.S. Structure-based engineering of E. coli galactokinase as a first step toward in vivo glycorandomization. *Chem. Biol.* **12**, 657-664 (2005).
2. Barton, W.A., Biggins, J.B., Jiang, J., Thorson, J.S. & Nikolov, D.B. Expanding pyrimidine diphosphosugar libraries via structure-based nucleotidylyltransferase engineering. *Proc. Natl. Acad. Sci. USA* **99**, 13397-13402 (2002).

3. Williams, G.J., Zhang, C. & Thorson, J.S. Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. *Nat. Chem. Biol.* **3**, 657-662 (2007).

4. Williams, G.J. & Thorson, J.S. A high-throughput fluorescence-based glycosyltransferase screen and its application in directed evolution. *Nat. Protocols* **3**, 357-362 (2008).

5. Williams, G.J., Goff, R.D., Zhang, C. & Thorson, J.S. Optimizing glycosyltransferase specificity via "hot spot" saturation mutagenesis presents a catalyst for novobiocin glycorandomization. *Chem. Biol.* **15**, 393-401 (2008).

6. Jiang, J., Albermann, C. & Thorson, J.S. Application of the nucleotidylyltransferase Ep toward the chemoenzymatic synthesis of dTDP-desosamine analogues. *Chembiochem* **4**, 443-446 (2003).

7. Fu, X. et al. Antibiotic optimization via in *vitro* glycorandomization. *Nat. Biotechnol.* **21**, 1467-1469 (2003).