

Supporting Information

Processive Carbohydrate Polymerases Mediate Bifunctional Catalysis Using a Single Active Site

John F. May, Matthew R. Levengood, Rebecca A. Splain, & Laura L. Kiessling

Table S1 GlfT2 homologs used in the multiple sequence alignment.

Abbreviation	GenBank Accession Number or NCBI Reference Sequence	Source of sequence (bacterial species)
Nbr	BAF46769.1	<i>Nocardia brasiliensis</i>
Nf1	YP_116420.1	<i>Nocardia farcinica</i> IFM 10152
Rr1	YP_704064.1	<i>Rhodococcus</i> sp. RHA1
Ser	YP_001107843.1	<i>Saccharopolyspora erythraea</i> NRRL 2338
Bde	ZP_02917795.1	<i>Bifidobacterium dentium</i> ATCC 27678
Ban	ZP_02964173.1	<i>Bifidobacterium animalis</i> subsp. lactis HN019
Bll	NP_695433.1	<i>Bifidobacterium longum</i> NCC2705
Mth	NP_218325.1	<i>Mycobacterium tuberculosis</i> H37Rv
Mtc	NP_338467.1	<i>Mycobacterium tuberculosis</i> CDC1551
Mma	YP_001853631.1	<i>Mycobacterium marinum</i> M
Mul	YP_908340.1	<i>Mycobacterium ulcerans</i> Agy99
Mle	NP_301191.1	<i>Mycobacterium leprae</i> TN
Mav	YP_879503.1	<i>Mycobacterium avium</i> 104
Map	NP_959146.1	<i>Mycobacterium avium</i> subsp. paratuberculosis K-10
Mjl	YP_001073654.1	<i>Mycobacterium</i> sp. JLS
Mmc	YP_642179.1	<i>Mycobacterium</i> sp. MCS
Mva	YP_956420.1	<i>Mycobacterium vanbaalenii</i> PYR-1
Mgi	YP_001132430.1	<i>Mycobacterium gilvum</i> PYR-GCK
Msm	YP_890616.1	<i>Mycobacterium smegmatis</i> str. MC2 155
Mab	YP_001700925.1	<i>Mycobacterium abscessus</i>
Rr2	YP_704009.1	<i>Rhodococcus</i> sp. RHA1
Nf2	YP_116383.1	<i>Nocardia farcinica</i> IFM 10152
Cur	YP_001799546.1	<i>Corynebacterium urealyticum</i> DSM 7109
Cje	YP_249900.1	<i>Corynebacterium jeikeium</i> K411
Cg1	NP_602073.1	<i>Corynebacterium glutamicum</i> ATCC 13032
Cg2	YP_227122.1	<i>Corynebacterium glutamicum</i> ATCC 13032
Cg3	BAC00276.1	<i>Corynebacterium glutamicum</i> ATCC 13032
Cgr	YP_001139691.1	<i>Corynebacterium glutamicum</i> R
Cef	NP_739324.1	<i>Corynebacterium efficiens</i> YS-314
Cdi	NP_940506.1	<i>Corynebacterium diphtheriae</i> NCTC 13129
Se1	YP_001107844.1	<i>Saccharopolyspora erythraea</i> NRRL 2338

Se2	YP_001102444.1	<i>Saccharopolyspora erythraea</i> NRRL 2338
Rsa	YP_001624767.1	<i>Renibacterium salmoninarum</i> ATCC 33209
Afb	YP_832172.1	<i>Arthrobacter</i> sp. FB24
Aau	YP_948391.1	<i>Arthrobacter aureescens</i> TC1
Kra	YP_001363614.1	<i>Kineococcus radiotolerans</i> SRS30216
Krh	YP_001854751.1	<i>Kocuria rhizophila</i> DC2201
Njs	YP_922603.1	<i>Nocardioides</i> sp. JS614
Cms	YP_001709406.1	<i>Clavibacter michiganensis</i> subsp. <i>Sepedonicus</i>
Cmm	YP_001221762.1	<i>Clavibacter michiganensis</i> subsp. <i>michiganensis</i> NCPPB 382
Bl2	NP_695432.1	<i>Bifidobacterium longum</i> NCC2705
Sen	ZP_02701557.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Newport str. SL317
Seh	ZP_02665507.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Heidelberg str. SL486
Sew	ZP_02832148.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Weltevreden str. HI_N05-537
Sep	YP_001589014.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Paratyphi B str. SPB7
Ses	ZP_02663925.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Schwarzengrund str. SL480
Sev	ZP_02706803.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Virchow str. SL491
Sej	ZP_02653520.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Javiana str. GA_MM04042433
Rba	ZP_01014392.1	<i>Rhodobacterales bacterium</i> HTCC2654
Gox	YP_191889.1	<i>Gluconobacter oxydans</i> 621H

Nbr	256	GGFTRGLYEVSAAN-----EHADVILMDDDILCEPETVVRLNAFANMTV-EPTLV	304
Nf1	254	GGFTRGLYEVSAVN-----EHADVILMDDDILCEPETVLRRLNAFANLTV-EPTLV	302
Rr1	250	GGFTRGMYEVSGIN-----EHANLILMDDDILCEPESILRMNAFANVTT-EPTLV	298
Ser	272	GGFSRGLAESVRCR-----GRVNTLLI DDDVRLPEPETSRLMTFAACTA-EPSLV	320
Bde	298	GGFSRGMYEETLKAGA-----SSYTLLLLDDDAISEPEAILRSMQFADYTK-KPTIV	346
Ban	296	GGFSRGMYEETLKAGT-----SAHTLLLLDDDAISEPESII RAVQFADYCN-TPTIV	344
Bl1	259	GGFSRGMFETVKAGK-----SDYTLLLLDDDAISEPESILRAVHFADYTV-RSALV	307
Mth	234	GGYSRVMYEALKN-----TDCQQILFMDDDIRLEPDSILRVLAMHRFAK-APMLV	282
Mtc	234	GGYSRVMYEALKN-----TDCQQILFMDDDIRLEPDSILRVLAMHRFAK-APMLV	282
Mma	232	GGYSRVMYEALKN-----TDCQQILFMDDDIRLEPDSILRVLAMHRFAK-SPMLV	280
Mul	232	GGYSRVMYEALKN-----TDCQQILFMDDDIRLEPDSILRVLAMHRFAK-SPMLV	280
Mle	232	GGYSRVMYEALKN-----TDCQQILFMDDDIRLEPDSILRVLAMHRFAK-SPMLV	280
Mav	235	GGYSRVMYEALKN-----TDCQQILFMDDDIRLEPDSILRVLAMHRFAK-SPMLV	283
Map	229	GGYSRVMYEALKN-----TDCQQILFMDDDIRLEPDSILRVLALHRFAK-TPMLI	277
Mj1	232	GGYSRVMYEALKN-----TDCQQILFMDDDIRLEPDSILRVLALHRFAK-TPMLI	280
Mmc	245	GGYSRVMYEALKN-----TDCEQILFMDDDIRVEPDSVLRALALNRFK-SPMLV	293
Mva	245	GGYSRVMYEALKN-----TDCEQILFMDDDIRVEPDSVLRALALNRFK-SPMLV	293
Mgi	254	GGYSRVMYEALKN-----TDCEQILFMDDDIRLEPDSILRALALNRFK-VPTLV	302
Msm	247	GGYSRVMYEALKN-----TDCEQILFMDDDIRVEPDSILRALALNRFK-VPTLV	295
Mab	245	GGYSRVMYEALKN-----TDCEQILFMDDDIRLEPDSILRALAMNRFK-SPILV	293
Rr2	245	GGYSRVMYEALKN-----TDCEQILFMDDDIRLEPDSILRALALNRFK-SPMLV	293
Nf2	245	GGYSRVMYEALKN-----TDCEQILFMDDDIRLEPDSILRALALNRFK-SPMLV	293
Cur	246	GGYSRIMYEALKT-----TDCQHILFMDDDIRLEPDSILRALAMSRFAK-TPMLV	294
Cje	245	GGYSRVMYEALKT-----TDAEYIVFMDDDIRLEPDSILRALAFARFAK-SPVLV	293
Cg1	269	GGYSRIMYEAERHAEGR-----TSPFILYMDDDIRLEPDSVLRSLAAARYAR-SPMLI	320
Cg2	276	GGYSRIMYEALHEERGT-----TSPYILYMDDDIRLEPDSVLRALAAARYAK-SPMLI	327
Cg3	250	GGYSRIMFEALGGVDGKGEAGAAKSPYILYMDDDIRLEPDSVLRALQVARYAK-SPILV	307
Cgr	253	GGYSRIMFEALGGVDGKGEAGAAKSPYILYMDDDIRLEPDSVLRALQVARYAK-SPILV	310
Cef	243	GGYSRIMFEALGGVDGKGEAGAAKSPYILYMDDDIRLEPDSVLRALQVARYAK-SPILV	300
Cdi	243	GGYSRIMFEALGGVDGKGEAGAAKSPYILYMDDDIRLEPDSVLRALQVARYAK-SPILV	300
Se1	285	GGYSRIMFEALGGVDGTGEAGASKSPYILYMDDDIRLEPDSVLRALQAARYAK-SPMLI	342
Se2	242	GGYSRIMYEALGGVDGTPAGAAQSPYILYMDDDIRLEPDSVLRALQVARYAK-SPILV	299
Rsa	263	GGYARIMYEALTT-----TDCEQILFMDDDIRLEPDSVLRALAFSRFAK-QPMLV	311
Afb	260	GGYGRVMYEALVYN-----SDAEQVLLI DDDIALEPDGVLNANAFARASS-QPVIV	308
Aau	265	GGFSRGMLEAVDNG-----SDYVLLI DDDVMVEPESINRLITFAELCK-KPTIV	312
Kra	271	GGFARGMFEAVENG-----SDYVLLMDDDIRVEPESII RLLTFADRCK-TPTIV	318
Krh	271	GGFARMYEMVVS DK-----SDYVLLI DDDIELETEGVMRAVAFADLCR-KPTIV	319
Njs	268	GGFARGMFEAVRGER-----SKYVLLI DDDVNIPEGIVRAVQFGDACR-KPTLV	316
Cms	236	GGFARGMYEASREGG-----SRYVLLI DDDVEVHPESILRAVRFGDFAR-HNTIV	284
Cmm	264	GGYARGQLESVRKGT-----ATYTMMDDIRVEPEGEVIRAITFADLAR-RPTIV	312
Bl2	271	GGFSRGMYEETLKEGA-----SDYVLMDDDIRLEPESIRRAVKFADYAR-TPTIV	319
Sen	271	GGFSRGMYEETLKEGA-----SDYVLMDDDIRLEPESIRRAVKFADYAR-TPTIV	319
Seh	271	GGFSRGMYEETLKEGA-----SDYVLMDDDIRLEPESIRRAVKFADYAR-TPTIV	319
Sew	248	GGFARGMYEVEHGE-----SGYALLI DDDTVLEPESVSRAIAFANHCE-KPTLV	296
Sep	213	GGFMRGLIEAGKIN-----DIKHVIFMDDDIRGSCIEIESICRTHAFLLMAKDKN TVV	262
Ses	213	GGFMRGLIEAGKIN-----DIKHVIFMDDDIRGSCIEIESICRTHAFLLMAKDKN TVV	262
Sev	213	GGFMRGLIEAGKIN-----DVKHVIFMDDDIRGSCIEIESICRTHAFLLMAKDKN TVV	262
Sej	213	GGFMRGLIEAGKIN-----DVKHVIFMDDDIRGSCIEIESICRTHAFLLMAKDKN TVV	262
Rba	213	GGFMRGLIEAGKIN-----DVKHVIFMDDDIRGSCIEIESICRTHAFLLMAKDKN TVV	262
Gox	213	GGFMRGLIEAGKIN-----DVKHVIFMDDDIRGSCIEIESICRTHAFLLMAKDKN TVV	262
	213	GGFMRGLIEAGKIN-----DIKHVIFMDDDIRGSCIEIESICRTHAFLLMAKDKN TVV	262
	213	GGFMRGLIEAGKIN-----DVKHVIFMDDDIRGSCIEIESICRTHAFLLMAKDKN TVV	262
	208	GGFTRTMVEATSVATP-----ATHHLLMDDDIRILDPRVNRALGFLAYVE-GELAV	257
	169	AGFTRGII EALSDPN-----ITHVVL MDDDIRVEVDAGLLCRIR SALAYIS-POICI	217

Figure S1. DDD motif is conserved among GlfT2 variants. Sequences are labeled according to Table S1, and the label for GlfT2 is in red. The DDD motif (Asp256–Asp258 for GlfT2) is invariant among these homologs, and the nearby residues are conserved. This alignment is colored using ClustalW2 coloring to indicate conservation.

Nbr	383	WDDVEYGI RAREAGFVTVTL PNAAVWHADFYWKDYD--DWARYFSMRNSLIVGALHTDLD	440
Nf1	381	WDDVEYGL RAREHG FVTVTL PNAAVWHADFYWKDYD--DWARYFSTRNSLIVGAMHTDLD	438
Rr1	377	WDDIEYGI RARAAGFVTVTL PNAGVWHADFWHKDRD--DWAKYFSIRNSLIAAALHSDFD	434
Ser	414	WDDVEFGL RARAAGYPTIAL PGAGVWHGDSRWRSDQ--DHAGYFHLRNGLITAGLHGGFG	471
Bde	430	FDDTEYAVRALEHG YHTVCL PGVAVVWHQAWHDKDP SRTWEEYFFQRNRWICGLLHCPKP	488
Ban	428	FDDIEYGL RAKEHG FHTVSL PGVAVVWHQAWHDKDP GRTWEEYFNRRNRWVCALLHAEKP	486
Bl1	391	FDDIEYGL RAKEHG VPTISL PGVAVVWHMGWHDKDP ARGWEEYFTQRNRWICALLHFPNA	449
Mth	370	WDDADYGL RAAEHG YPTVTL PGAAIWHMAWS DKDDAIDWQAYFHLRNRLVVAAMHWDGP	428
Mtc	370	WDDADYGL RAAEHG YPTVTL PGAAIWHMAWS DKDDAIDWQAYFHLRNRLVVAAMHWDGP	428
Mma	370	WDDADYGL RAAEHG YPTVTL PGAAIWHMAWS DKDDAIDWQAYFHLRNRLVVAAMHWDGP	428
Mul	368	WDDADYGL RAAEHG YPTVTL PGAAIWHMAWS DKDDAIDWQAYFHLRNRLVVAAMHWDG-	425
Mle	368	WDDADYGL RAAEHG YPTVTL PGAAIWHMAWS DKDDAIDWQAYFHLRNRLVVAAMHWDG-	425
Mav	371	WDDADYGL RAAEHG YPTVTL PGAAIWHMAWS DKDDAIDWQAYFHLRNRLVVAAMHWDG-	428
Map	365	WDDAEYGL RAGEHG YPTVTL PGAAIWHMAWS DKDDAIDWQAYFHLRNRLVVAALHWDG-	422
Mj1	368	WDDAEYGL RAGEHG YPTVTL PGAAIWHMAWS DKDDAIDWQAYFHLRNRLVVAALHWDG-	425
Mmc	381	WDDAEYGL RAAEQGYPTATMPGTAIWHMAWS DKDDAIDWQAYFHLRNRLVVAALHWDG-	438
Mva	381	WDDAEYGL RAAEQGYPTATMPGTAIWHMAWS DKDDAIDWQAYFHLRNRLVVAALHWDG-	438
Mgi	391	WDDADYGL RAGEHG YPTVTL PGAAIWHMAWS DKDDAIDWQAYFHLRNRLVVAALHWDG-	448
Msm	384	WDDAEYGL RAGEHG YPTVTL PGAAIWHMAWS DKDDAIDWQAYFHLRNRLVVAALHWDG-	441
Mab	382	WDDAEYGL RAGEHG YGTVTL PGAAIWHMAWS DKDDAIDWQAYFHLRNRLVVAALHWDG-	439
Rr2	382	WDDAEYGL RANEHG YGTASMPGTAIWHMAWS DKDDAIDWQAYFHLRNRLVVAALHWDN-	439
Nf2	382	WDDAEYGL RANEHG YGTASMPGTAIWHMAWS DKDDAIDWQAYFHLRNRLVVAALHWDN-	439
Cur	381	WDDAEYGL RARDAGYPTVTMPGTAIWHMAWS DKDDAIDWQAYFHLRNRLVVAALHMPG-	438
Cje	380	WDDVEYGL RAREAGYPTVTL PGAAVWHMAWS DKDDAIDWQAYFHLRNRLVVAALHLPG-	437
Cg1	418	WDDGEFGL RAKDAGFP TASWPGIAIWHMAWS DKDDAIDWQAYFHLRNRLVVAALHHDG-	475
Cg2	445	WDDGEFGL RAKDAGFP TASWPGIAIWHMAWS DKDDAIDWQAYFHLRNRLVVAALHHDG-	502
Cg3	403	WDDAEYGL RARKAGFP TATWPGIAIWHMAWS DKDDAIDWQAYFHLRNRLVVAAMYHOG-	460
Cgr	406	WDDAEYGL RARKAGFP TATWPGIAIWHMAWS DKDDAIDWQAYFHLRNRLVVAAMYHOG-	463
Cef	396	WDDAEYGL RARKAGFP TATWPGIAIWHMAWS DKDDAIDWQAYFHLRNRLVVAAMYHOG-	453
Cdi	396	WDDAEYGL RARKAGFP TATWPGIAIWHMAWS DKDDAIDWQAYFHLRNRLVVAAMYHOG-	453
Se1	438	WDDAEYGL RAGAAGFP TATWPGIAIWHMAWS DKDDAIDWQAYFHLRNRLVVAAMNHDG-	495
Se2	395	WDDAEYGL RAGNAGFATATWPGVAIWHLAWSDKDDAIDWQAYFHLRNRLVVAALHGGP-	451
Rsa	395	WDDVEYGL RAGAAGFP TVTVPGIAVWHMSFADKHD SLNQWEYFLTRNQLVVAALHGGP-	452
Afb	392	WDDAEYSL RAGGHG YPTVTL PGAAIWHMPWTDKNDATDWTAYFHTRNRLILAALHSPDE	450
Aau	397	WDDAEYGL RAKQHG YATVSL PGAAVWHVSWIDKDDL VGWQAYFHTRNRLITALLHSPYE	455
Kra	403	WDDSEYGL RAKAHG FPTVSL PGS AVWHVSWIDKDDL VGWQAYFHARNRVIAALLHSPYE	461
Krh	402	WDDAEYSL RGRAAGFP TVTL PGAAVWHVSWADKDDSDVDWQAYYHERNRLIATLLHSPFP	460
Njs	400	WDDAEYGL RAMEHG YSTVSL PGACVWHVSWNDKDDTIDWQAYYHERNRLVALLHSPYE	458
Cms	368	WDDAEFGL RAKQAG YSTVSL PGVAVVWHMPWTEKDDTIDWQAYYHARNRWLAALLHSPYS	426
Cmm	396	WDDSEFGL RAKEAG YPTVTF PGAAVWHVPWTDKNDGLDWQAYFHQRNRFVAALLHSPYP	454
Bl2	401	WDDAEYAL RAKEVG VPTVTL PGAAVWHVSWVDKDDSDQWQAF FHARNRLIAALLHSPYE	459
Sen	401	WDDAEYAL RAKEVG VPTVTL PGAAVWHVSWVDKDDSDQWQAF FHARNRLIAALLHSPYE	459
Seh	380	NDDVEYGVRAQRAGYRTVTVPGVCLWHQS FVDKDDQLDWQAYYHIRNRTIMGLLYANQQ	438
Sew	338	GDDLDFGYMHKKH--NIVTLNGVASWQMDFER--KISVLNSYLNFRFVAVPALISKRK-	391
Sep	338	GDDLDFGYMHKKH--NIVTLNGVASWQMDFER--KISVLNSYLNFRFVAVPALISKRK-	391
Sev	338	GDDLDFGYMHKKH--NIVTLNGVASWQMDFER--KISVLNSYLNFRFVAVPALISKRK-	391
Sej	338	GDDLDFGYMHKKH--NIVTLNGVASWQMDFER--KISVLNSYLNFRFVAVPALISKRK-	391
Rba	338	GDDLDFGYMHKKH--NIVTLNGVASWQMDFER--KISVLNSYLNFRFVAVPALISKRK-	391
Gox	338	GDDLDFGYMHKKH--NIVTLNGVASWQMDFER--KISVLNSYLNFRFVAVPALISKRK-	391
	338	GDDLDFGYMHKKH--NIVTLNGVASWQMDFER--KISVLNSYLNFRFVAVPALISKRK-	391
	338	GDDLDFGYMHKKH--NIVTLNGVASWQMDFER--KISVLNSYLNFRFVAVPALISKRK-	391
	338	GDDLDFGYMHKKH--NIVTLNGVASWQMDFER--KISVLNSYLNFRFVAVPALISKRK-	391
	335	GDDIEYGCMAANGVETICLPGVAVWHESFHHK--TSDWLTYYDMRNRLFVASLYPOLV	391
	298	GDDAEYGLRLKRAGFPTVMWPGVYVAHPNLQN--QTRPWHHYDRRNALICALLERGAV	354

Figure S2. DDA motif is conserved among GlfT2 variants. Sequences are labeled according to Table S1, and the label for GlfT2 is in red. The Asp residues in the DDA motif (Asp371–Ala373 for GlfT2) are invariant among these homologs, and nearby residues are conserved. This alignment is colored using ClustalW2 coloring to indicate conservation.

```

Glft2 160 -----
2z86A 111 PLDWPSDLTL PPLPESTNDY VWAGKRKELL II--DGLSIV IPTYNRAKIL

Glft2 176 VNALRELTAD -PLVDQVIGA VIVPDQGERK VRDHPDFPAA AARLGSRLSI
2z86A 166 AITLACLNCQ KTIYD---YE VIVADDGSK- --EN--IEEI VREFESLLNI

Glft2 225 HDQPNLGGSG GYSRVMYEAL KNTDCQQILF MDDDIRLEPD SILRVLAMHR
2z86A 208 KYVRQKDYGY QLC AVRNLGL RAAKYNVVAI LDCDMAPNPL WVQSYMELLA

Glft2 275 FAKAPMLVGG QMLNLQE-PS HLHIMGEVVD RSIFMWTAAP HAEYDHDFAE
2z86A 258 VDDNVALIGP RKYIDTSKHT Y----LDFLS QKSLINEIPE SV--DWRIEH

Glft2 324 YPLNDNNSRS KLLHRRIDVD YNGWWTMIP RQVAEELGQP LPLFIKWDDA
2z86A 317 FKN-TDN--- -LRLCNTPFR FFSGGNVAFA KKWLFragWF DEEFTHWGGE

Glft2 374 --DYGLRAAE HGYPTVTLPG AAIWHMA--- -----
2z86A 362 DNEFGYRLYR EGCYFRSVEG AMAYHQEPPQ LLQQKVPYFY RKKEKIESAT

```

Figure S3. Alignment used for homology model of Glft2. This alignment was generated by the SWISS-MODEL server and used as the basis for the homology model shown in Figure 3B. This region of the Glft2 sequence corresponds to its GT2 domain. The DDD and DDA motifs in Glft2 are highlighted in red. This alignment showed 14.8% identity between the two sequences with an e-value of 3.6×10^{-9} .

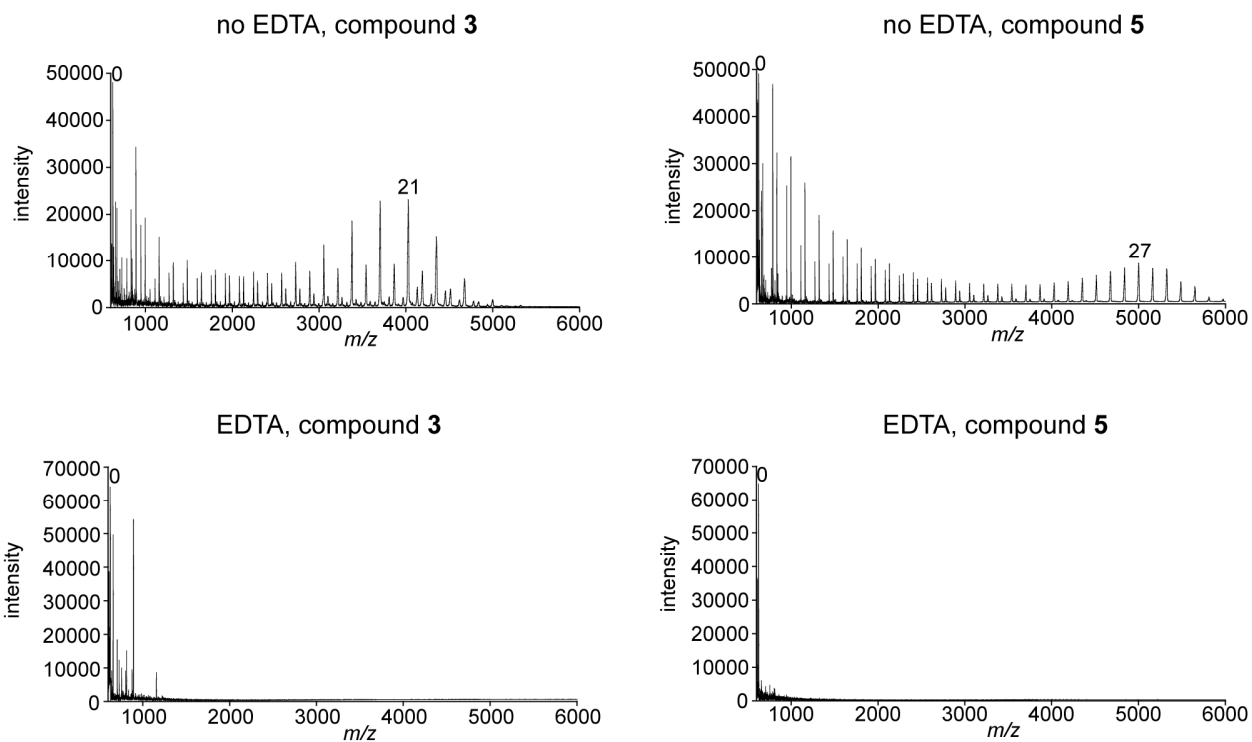


Figure S4. GlfT2 catalysis is metal-dependent. Spectra from MALDI-TOF MS analysis of GlfT2 reactions are shown. Products corresponding to addition of n Gal f residues to compound **3** (left) or compound **5** (right) are observed from reactions of GlfT2, acceptor, and UDP-Gal f only when metal (Mg^{2+} ; “no EDTA”) is present. In the presence of EDTA and absence of added metal, no products are observed from reactions with either acceptor (bottom spectra).

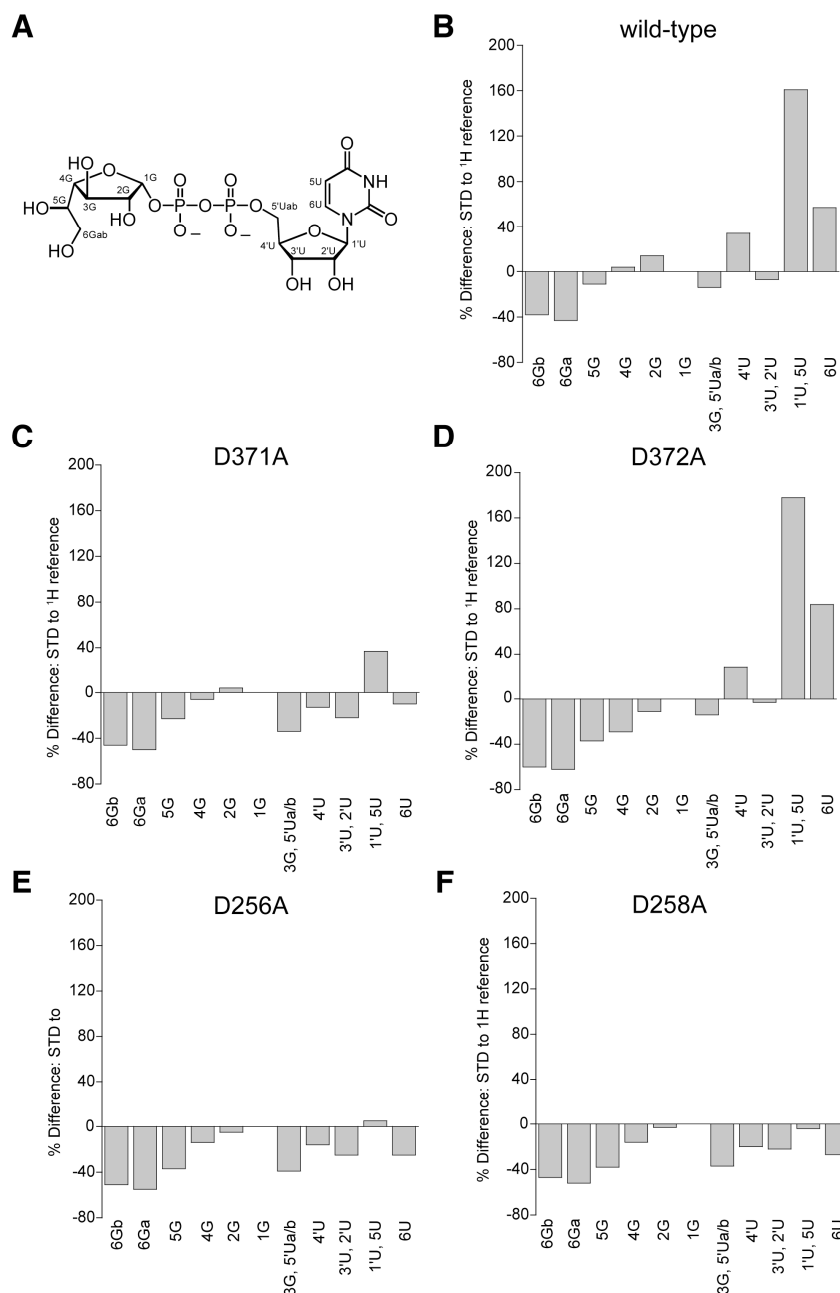


Figure S5. Analysis of STD NMR spectra for GlfT2 variants binding to UDP-Galf. Addition of wild-type GlfT2 (B), D371A GlfT2 (C), D372A GlfT2 (D), D256A GlfT2 (E), or D258A GlfT2 (F) to a solution containing UDP-Galf results in saturation transfer difference NMR signals at the indicated protons in the UDP-Galf ligand (A). The signal for each distinct proton is plotted to give an STD NMR profile. Comparison of STD NMR profiles for the variants indicates that D372A and wild-type GlfT2 bind UDP-Galf in a similar manner, D256A and D258A bind UDP-Galf similarly to each other but distinctly from wild-type GlfT2, and D371A binds UDP-Galf in a manner distinct from wild-type GlfT2.

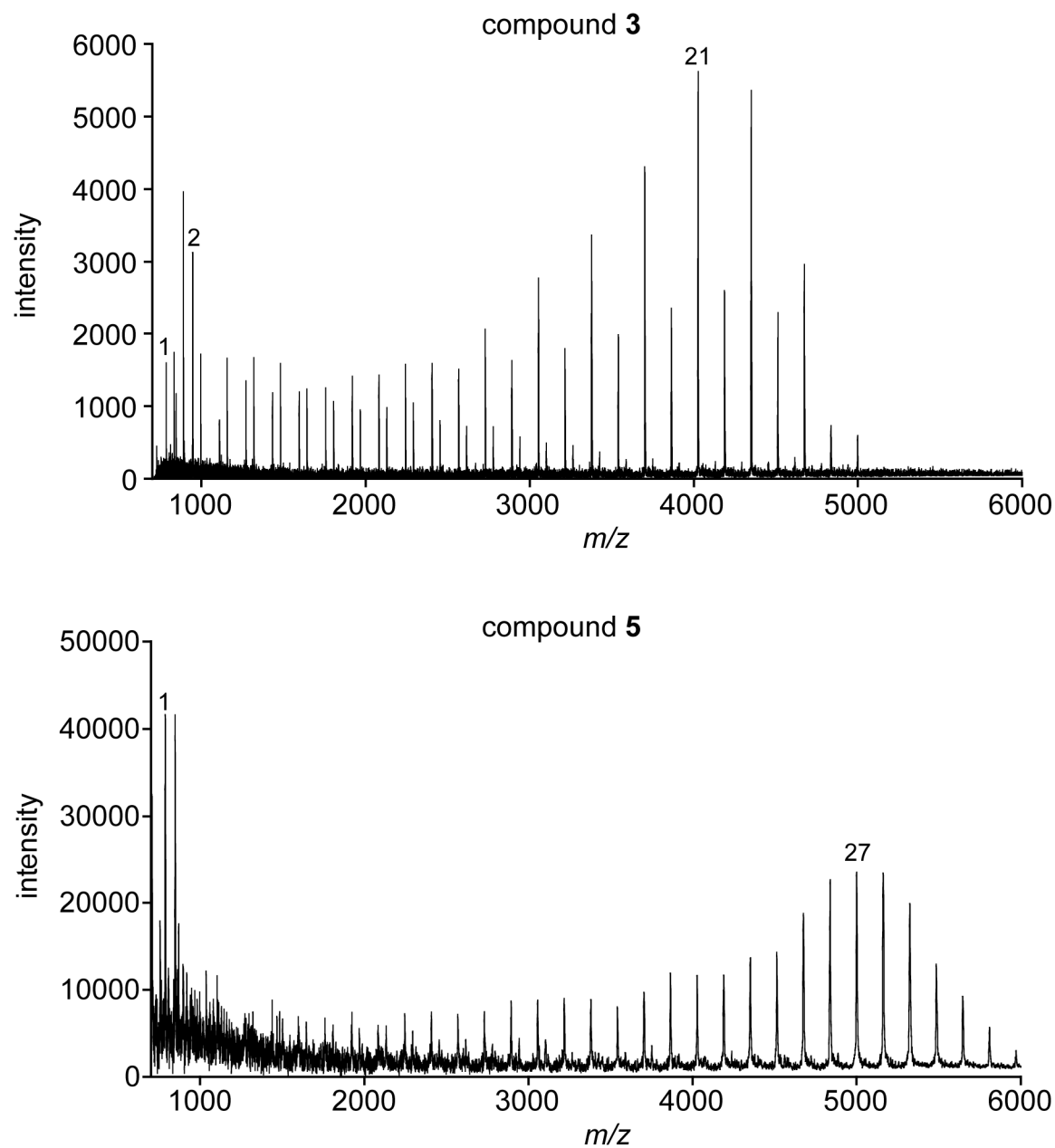


Figure S6. GlfT2 D371E variant shows wild-type polymerization activity. Spectra from MALDI-TOF MS analysis of 20 h incubations of GlfT2 D371E with either compound **3** (top) or compound **5** (bottom) show polymeric products. To orient these spectra with those shown in Figure 5, certain peaks that correspond to $[M+Na]^+$, where M equals the mass of the acceptor plus n Gal f residues, are labeled with the appropriate value of n . Labels are shown only for the short products ($n = 1, 2$) and for the highest-intensity peak of the polymeric products ($n = 21, 27$).

A

```

SpHasA  41  -----KVAAV  IPSYNEDAES  LLETLKSVLA  QTYPLSEIYI  VDDGSSNTDA
3bcvB   3  LIP--KSVVI  VPIYNVE-KY  LDQCQVQALLA  QTLSDIEIIL  IDDESP-DNC

SpHasA  86  IQLIEEYVNR  EVDICRNVIV  HRSLVNKGKR  HAQAWAFERS  DADVFLTVDS
3bcvB   49  PKICDDYAAQ  YPN-----IK  VIHKKNAGLG  XACNSGLDVA  TGEYVAFCDL

SpHasA 136  DTYIYPNALE  ELLKSFNDET  VYAATGHLNA  RNRQTNLLTR  LTDIRYDNAF
3bcvB   94  DDYVDSXYYX  TXYNVAQKYT  CDAVFTFKLY  KNKNEI-HTL  LKDLIASDPY

SpHasA 186  GVERAAQSLT  GNILVCSGPL  SIYRREVIIP  NLERYKNQTF  LGLPVSIGDD
3bcvB  163  AREE-----  -RAIQVSAKV  VLYRRNLIEK  KHLRFVSER-  ----ILPSED

SpHasA 236  RCLTNYAIDL  GRTVY  --
3bcvB  201  LIFNVDVLAN  SNIVCVLP

```

B

```

CesA   145  -----PDE  WPTVDIFVPT  YNEELSIVRL  TVLGSLGIDW
1xhbA  95  NRSLPDVRL  GCKTKVYPDN  LPTTSVVIVF  HNEAWSTLLR  TVHSVINRSP

CesA   178  PPEKVRVHIL  DDGRRPE-FA  AFAAECG---  ---ANYIARP  TNEHAKAGNL
1xhbA  145  RHMIEEIVLV  DDASERDFLK  RPLESYVKKL  KVPVHVIRME  QRS-GLIRAR

CesA   221  NYAIGHTDGD  YILIFDCDHV  PTRAFQLQTM  GWMVEDPKIA  LMQTPHHFYS
1xhbA  194  LKGAAVSRGQ  VITFLDAHCE  CTAGWLEPLL  ARIKHDRRTV  VCPPII-DVIS

CesA   271  PDPFQRNLSA  GYR-----  -----TPPE  GNLFGYGVVQD  GNDFWDATFF
1xhbA  243  DDTFEYMAGS  DMTYGGFNWK  LNFRWYPVPQ  REMDRR-KGD  RTLPVRTPTM

CesA   308  CGSCAILRRT  AIEQIGGFAT  QTV---TEDA  HTALKMQRLG  WSTAYL  ---
1xhbA  292  AGGLFSIDRD  YFQEIGTYDA  GMDIWGGENL  EISFRIWQCG  GTLEIVTCSH

```

Figure S7. Sequence alignments used for homology modeling of other GT-2 glycosyltransferases. (A) Alignment of hyaluronan synthase from *Streptococcus pyogenes* (SpHasA; GenBank accession number AAA17984.1) with the sequence of a glycosyltransferase PDB ID: 3BCV. This alignment, generated automatically by SWISS-MODEL, was used for the subsequent homology model of SpHasA, which showed 18.6% identity between the two sequences and gave an e-value of 1.3×10^{-14} . (B) Alignment of cellulose synthase from *Gluconacetobacter xylinus* (CesA; GenBank accession number P21877.3) with the glycosyltransferase UDP-GalNAc:polypeptide α -N-acetylgalactosaminyltransferase-T1 (PDB ID: 1XHB). This alignment was generated automatically in SWISS-MODEL and was used for subsequent homology modeling. This alignment showed 13.1% sequence identity between the two sequences with an e-value of 1.6×10^{-11} .

Table S2. Oligonucleotide primers used for site-directed mutagenesis. For each desired mutation, the primer indicated in the table and its reverse complement (not shown) were used together in the PCR mixture.

GlfT2 D256A	5'— CAGATCCTATTCATGG <u>GCC</u> GACGACATCCGCCTC—3'
GlfT2 D256E	5'— CAGATCCTATTCATGG <u>GAG</u> GACGACATCCGCCTC—3'
GlfT2 D258A	5'— CTATTCATGGACGACG <u>GCC</u> ATCCGCCTCGAGCCG—3'
GlfT2 D258E	5'— CTATTCATGGACGACG <u>GAG</u> ATCCGCCTCGAGCCG—3'
GlfT2 D371A	5'— GCCGTTGTTTCATCAAATGGG <u>GCC</u> GACGCCGATTACGGCC—3'
GlfT2 D371E	5'— GCCGTTGTTTCATCAAATGGG <u>GAA</u> GACGCCGATTACGGCC—3'
GlfT2 D372A	5'— GCCGTTGTTTCATCAAATGGGACG <u>GCC</u> GCCGATTACGGCC—3'
GlfT2 D372E	5'— GCCGTTGTTTCATCAAATGGGACG <u>GAA</u> GCCGATTACGGCC—3'