**Supplementary Information**

# Notch-modifying xylosyltransferase structures support an $S_Ni$-like retaining mechanism

HongjunYu[1], Megumi Takeuchi[2,#], Jamie LeBarron[3], Joshua Kantharia[2], Erwin London[2], Hans Bakker[4], Robert S. Haltiwanger[2,#], Huilin Li[1,2,*], Hideyuki Takeuchi[2,*,#]

[1] Biosciences Department, Brookhaven National Laboratory, Upton, NY 11973, USA.

[2]Department of Biochemistry and Cell Biology, Stony Brook University, Stony Brook, NY 11794, USA.

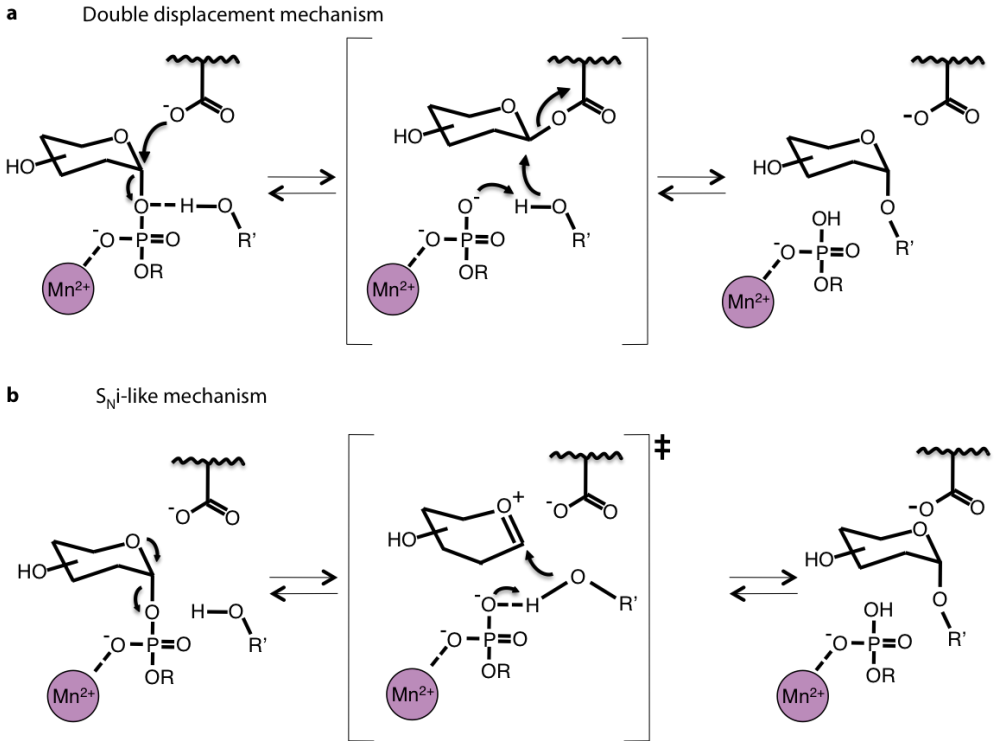
[3]Department of Physiology and Biophysics, Stony Brook University, Stony Brook, NY 11794, USA

[4] Department of Cellular Chemistry, Hannover Medical School, Carl-Neuberg-Strasse 1, 30625 Hannover, Germany.

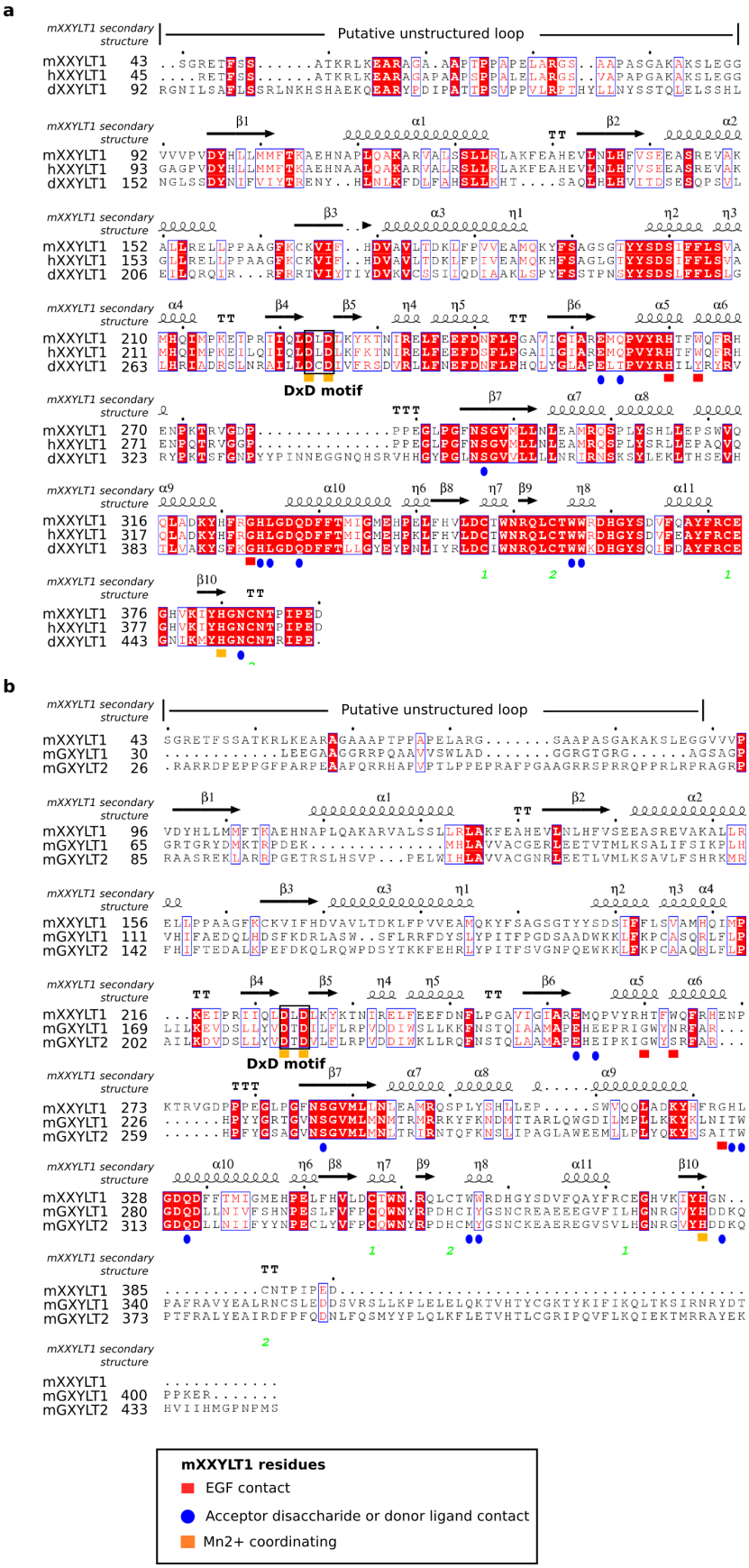
#Present address: Complex Carbohydrate Research Center, The University of Georgia, Athens, GA 30602, USA

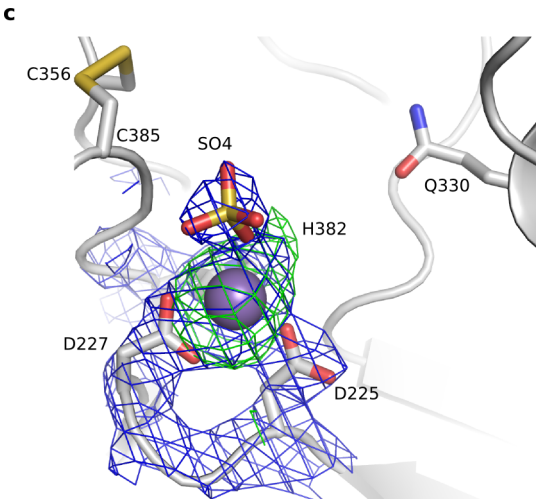
*Correspondence to: H.L. (hli@bnl.gov) or H.T. (takeuchi@uga.edu).

## Supplementary Results

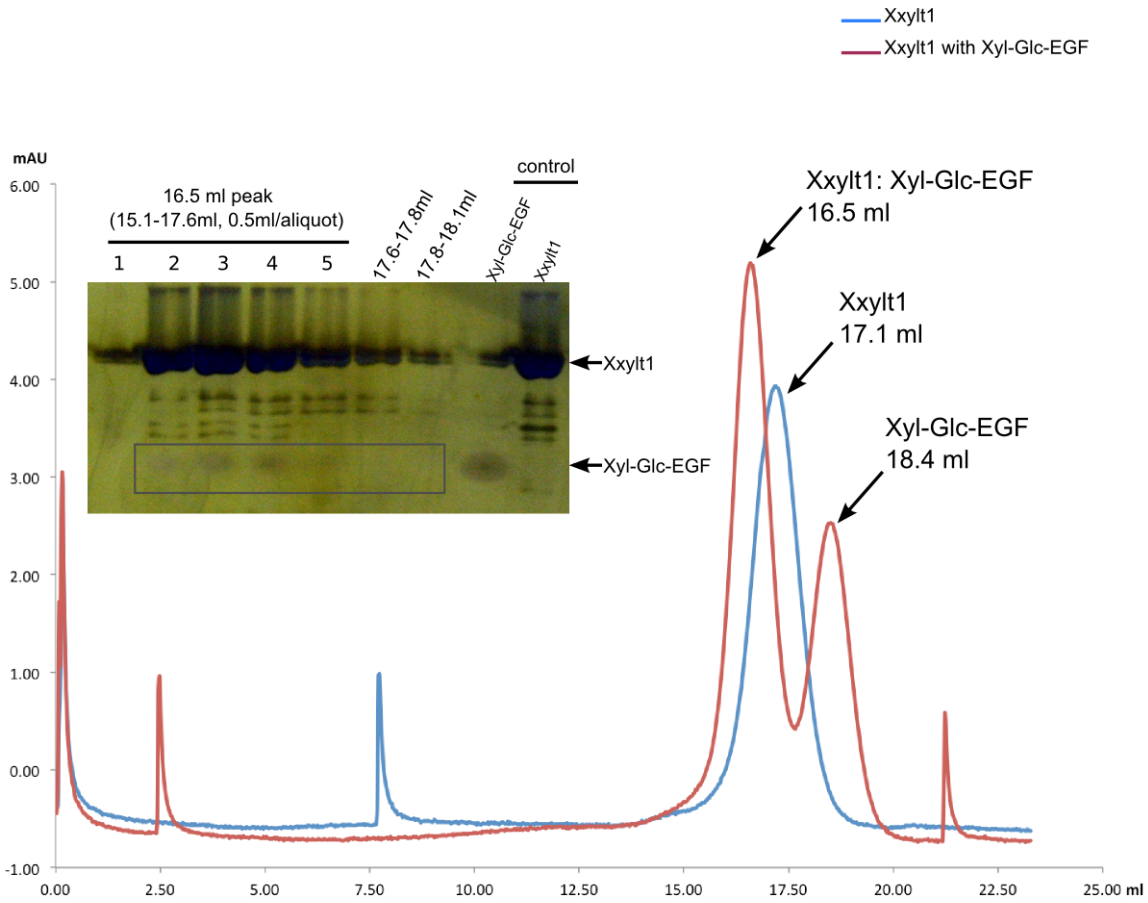



**Supplementary Figure 1. Two possible reaction mechanisms for retaining glycosyltransferase. (a)** Double displacement. This mechanism involves the formation of a covalently bound glycosyl-enzyme intermediate. **(b)** $S_N$i-like mechanism. $S_N$i is a form of $S_N$1 reaction in which the nucleophile is derived by decomposition of the leaving group and attacks from the same face. $S_N$i-like mechanism involves the formation of an oxocarbenium-like intermediate, followed by nucleophilic attack by the deprotonated hydroxyl from the acceptor toward the anomeric C1 atom of the donor sugar. R, a nucleoside monophosphate, and R'OH, an acceptor group.

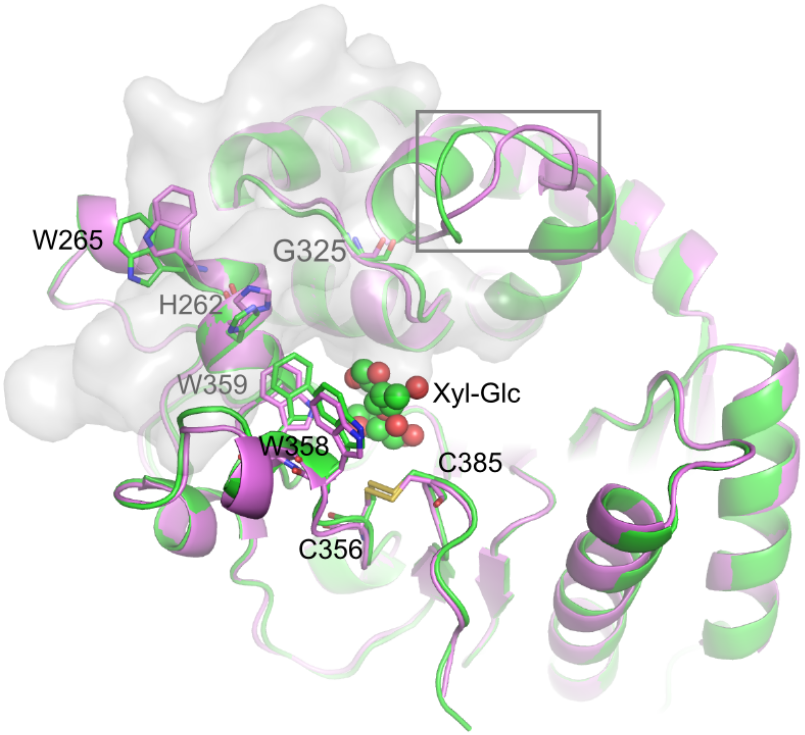

a
mXXYLT1 secondary structure
Putative unstructured loop
mXXYLT1 43 ..SGRETFSS.....ATKRLKEARAGA.AAPTPPAPELARGS..AAPASGAKAKSLEGG
hXXYLT1 45 ....RETFSS.....ATKRLKEARAGAPAAPSPPALELARGS..VAPAPGAKAKSLEGG
dXXYLT1 92 RGNILSAFLSSRLNKHSHAEKQEARYPDIPATTPSVPPVLRPTHYLLNYSSTQLELSSHL
mXXYLT1 92 VVVPVDYHLLMMFTKAEHNAPLQAKARVALSSLLRLAKFEAHEVLNLHFVSEEASREVAK
hXXYLT1 93 GAGPVDYHLLMMFTKAEHNAALQAKARVALRSLLRLAKFEAHEVLNLHFVSEEASREVAK
dXXYLT1 152 NGLSSDYNIFVIYTRENY..HLNLKFDLFAHSLLKHT....SAQLHLHVITDSESQPSVL
mXXYLT1 152 ALLRELLPPAAGFKCKVIF..HDVAVLTDKLFPVVEAMQKYFSAGSGTYYSDSIFFLSVA
hXXYLT1 153 GLLRELLPPAAGFKCKVIF..HDVAVLTDKLFPIVEAMQKHFSAGLGTYYSDSIFFLSVA
dXXYLT1 206 EILQRQIR...RFRTVIYTIYDVKVCSSILQDIAAKLSPYFSSTPNSYYSDSLFFLSLG
mXXYLT1 210 MHQIMPKETPRIIQLDLDLKYKTNIRELFEEFDNFLPGAVIGHARDMQPVYRHTFWQFRH
hXXYLT1 211 MHQLMPKEILQIIQLDLDLKFKTNIRELFEEFDSFLPGAIIGHARDMQPVYRHTFWQFRH
dXXYLT1 263 LHRRADRSLNRAILLDCDIVFRSDVRLLFNEFDNFLPHQLYGLAPELTPVYRHILYRYRV
DxD motif
mXXYLT1 270 ENPKTRVGDP.....PPEGLPCFNSGVMLLNLEANRQSPLYSHLLEPSWVQ
hXXYLT1 271 ENPQTRVGGP.....PPEGLPCFNSGVMLLNLEANRQSPLYSRLLEPAQVQ
dXXYLT1 323 RYPKNTSFGNPYYPINNEGGNQHSRVHHGYPCLNSGVLLLLNRIRNSKSYLEKLTHSEVH
mXXYLT1 316 QLADKYHFEGHLGDQDFFTMIGMEHPELFHVLDCTWNRQLCTWWRDHGYSDVFQAYFRCE
hXXYLT1 317 QLADKYHFEGHLGDQDFFTMIGMEHPKLFHVLDCTWNRQLCTWWRDHGYSDVFEAYFRCE
dXXYLT1 383 TLVAKYSFKGHLGDQDFFTLLGYEYPNLIYRLDCTWNRQLCTWWKDHGYSQIFDAYFRCE
mXXYLT1 376 GHVKIYHGNCNTPIPED
hXXYLT1 377 GHVKIYHGNCNTRIPED
dXXYLT1 443 GNIKMYHGNCNTRIPE.
b
mXXYLT1 43 SGRETFSSATKRLKEARAGAAAPTPPAPELARG.....SAAPASGAKAKSLEGGVVVP
mGXYLT1 30 ..........LEEGAAGGRRPQAAVVSWLAD.....GGRGTGRG.....AGSAGP
mGXYLT2 26 .RARRDPEPPGFPARPEAAPQRRHAPVPTLPPEPRAFPGAAGRRSPRRQPFRLRPRAGRP
mXXYLT1 96 VDYHLLMMFTKAEHNAPLQAKARVALSSLLRLAKFEAHEVLNLHFVSEEASREVAKALLR
mGXYLT1 65 GRTGRYDMKTRPDEK.....MHLAVVACGERLEETVTMLKSALIFSIKPLH
mGXYLT2 85 RAASREKLARRPGETRSLHSVP...PELWIHLAVVACGNRLEETLVMLKSAVLFSHRKMR
mXXYLT1 156 ELLPPAAGFKCKVIFHDVAVLTDKLFPVVEAMQKYFSAGSGTYYSDSIFFLSVAMHQIMP
mGXYLT1 111 VHIFAEDQLHDSFKDRLASW..SFLRRFDYSLYPITFPGDSAADWKKLFKPCASQRLFLP
mGXYLT2 142 FHIFTEDALKPEFDKQLRQWPDSYTKKFEHRLYPITFSVGNPQEWKKLFKPCAAQRLFLP
mXXYLT1 216 ...KEIPRIIQLDLDLKYKTNIRELFEEFDNFLPGAVIGHAREMQPVYRHTFWQFRHENP
mGXYLT1 169 LILKEVDSLLYVDTDILFLRPVDDIWSLLKKFNSTQIAAMAPEHEEPRIGWYNRFAR...
mGXYLT2 202 AILKDVDSLLYVDTDVLFLRPVDDIWKLLRQFNSTQLAAMAPEHEIPKIGWYSRFAR...
DxD motif
mXXYLT1 273 KTRVGDPPPFGLPGFNSGVMLLNLEAMRQSPLYSHILEP.....SWVQQLADKYFFRGHL
mGXYLT1 226 ......HPYYGRTGVNSGVMLMNLTRMRRKYFKNDMTTARLQWGDILMPLLKKYKLNITW
mGXYLT2 259 ......HPFYGSAGVNSGVMLMNLTRHRNTQFKNSLIPAGLAWEEMLLPLYQKYKSAITW
mXXYLT1 328 GDQDFFTMIGMEHPELFHVLDCTWN.RQLCTWWRDHGYSDVFQAYFRCEGHVKIYHGN..
mGXYLT1 280 GDQDLLNIVFSHNPESLFVFPCQWNYRPDHCIYGSNCREAEEEGVFILHGNRGVYHDDKQ
mGXYLT2 313 GDQDLLNIIFYYNPECLYVFPCQWNYRPDHCMYGSNCKEAEREGVSVLHGNRGVYHDDKQ
mXXYLT1 385 ..........CNTPIPED.....
mGXYLT1 340 PAFRAVYEALRNCSLEDDSVRSLLKPLELELQKTVHTYCGKTYKIFIKQLTKSIRNRYDT
mGXYLT2 373 PTFRALYEAIRDFPFQDNLFQSMYYPLQLKFLETVHTLCGRIPQVFLKQIEKTMRRAYEK
mXXYLT1 .....
mGXYLT1 400 PPKER.....
mGXYLT2 433 HVIIHMGPNPMS
mXXYLT1 residues
EGF contact
Acceptor disaccharide or donor ligand contact
Mn2+ coordinating



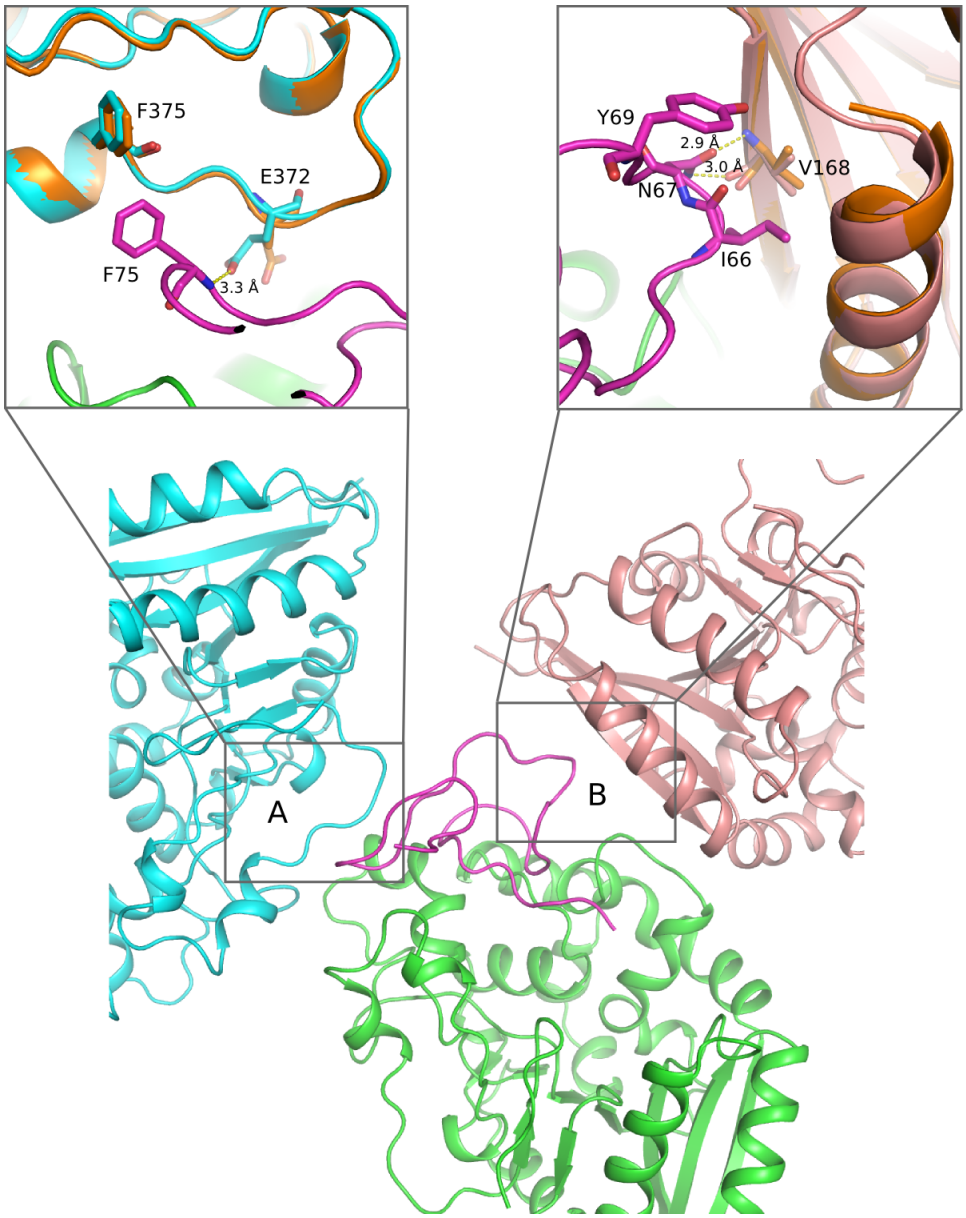

**Supplementary Figure 2. Sequence and structure features of XXYLT1.** (**a**) Sequence conservation of XXYLT1 extracellular domains among *Mus musculus* (mXXYLT1, accession number Q3U4G3), *Homo sapiens* (hXXYLT1, accession number Q8NBI6) and *Drosophila melanogaster* (dXXYLT1, accession number Q9W1D1). **(b)** Sequence conservation between *Mus musculus* XXYLT1, GXYLT1 and GXYLT2. (**c**) The DXD motif and the $Mn^{2+}$ binding site in XXYLT1. A nearby bound sulfate may mimic the binding of β-phosphate from donor ligand. The map was generated before modeling the $Mn^{2+}$ in the active site. The green mesh shows *Fo-Fc* difference density map contoured at 3.0σ, and the blue mesh shows *2Fo-Fc* map contoured at 1.8σ.



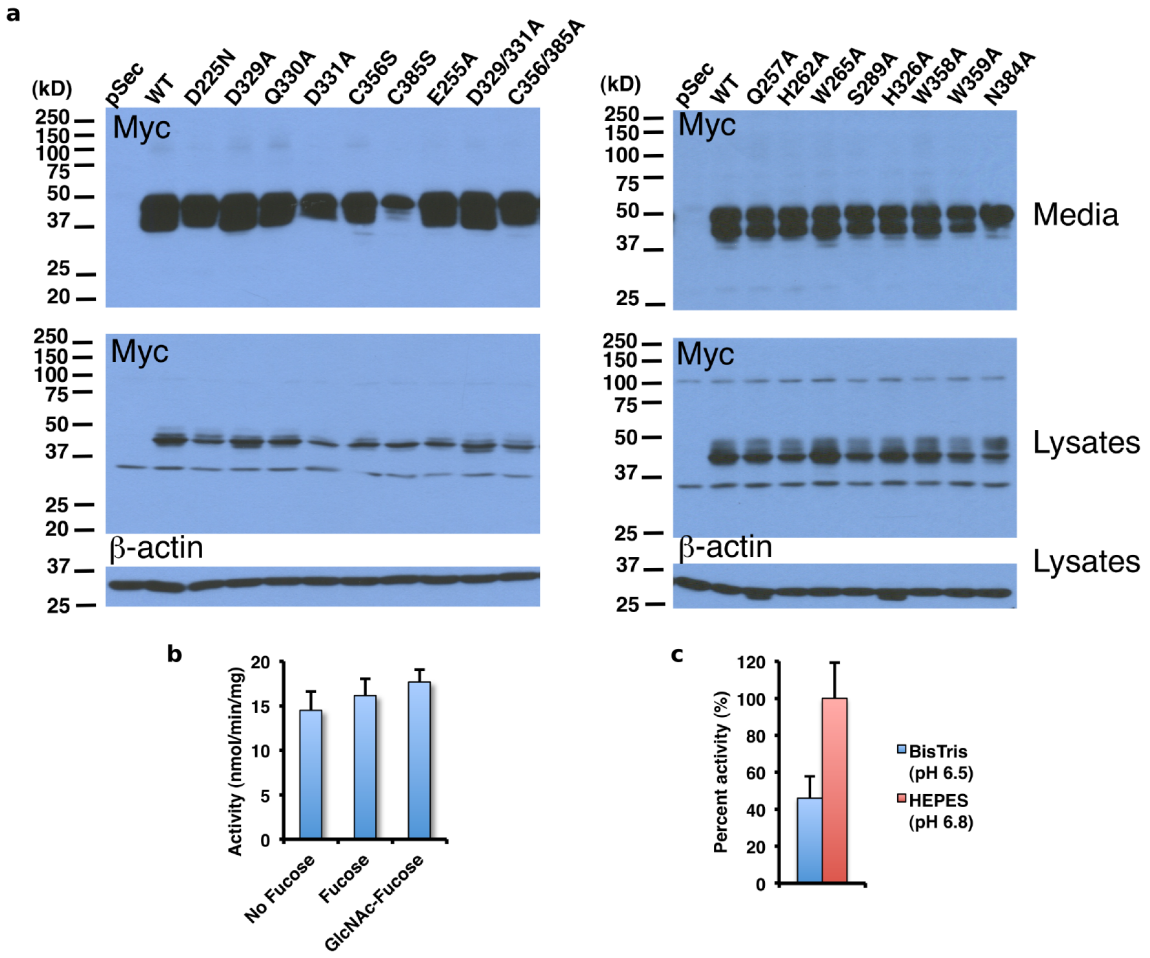

**Supplementary Figure 3. Gel filtration profile of the mixture of XXYLT1 and Xyl-Glc-EGF (red) compared with that of the purified XXYLT1 (blue).** One profile of two independent replicates was shown. Inset showed a silver-stained Tricine SDS-PAGE gel of the 16.5 ml peak of the mixed sample after scale-up sample preparation, showing the presence of both XXYLT1 and Xyl-Glc-EGF. Note that the three sharp peaks near 2.5 ml, 8 ml, and 21 ml were from instrumental noise due to the reciprocating pumps.



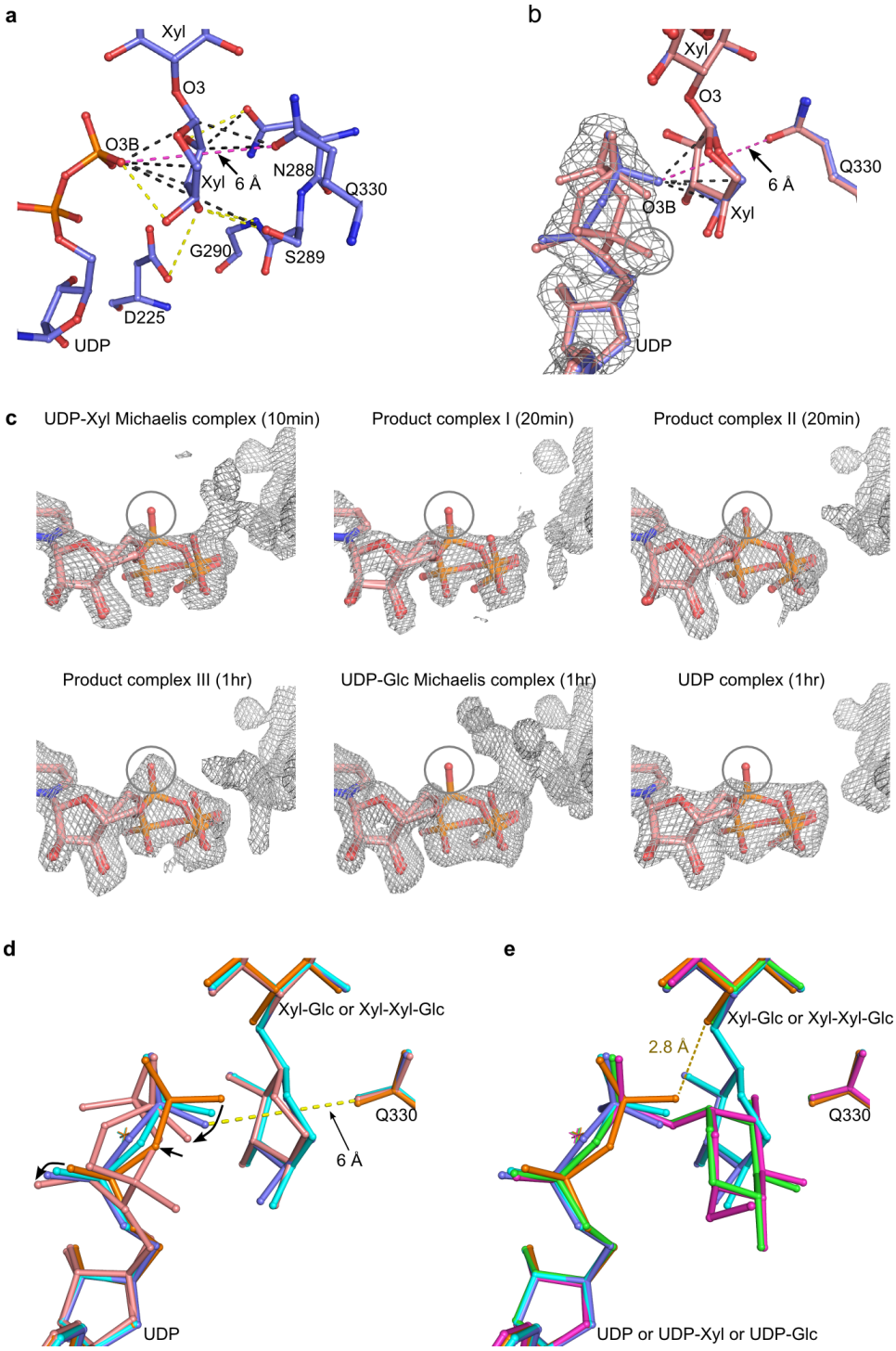

**Supplementary Figure 4. Superposition of the apo-XXYLT1 (in magenta cartoon) and the XXYLT1: Xyl-Glc-EGF binary complex structure (in green cartoon).** Only small conformational changes were detected in XXYLT1 structure upon Xyl-Glc-EGF binding, indicating that the EGF-interacting interface in XXYLT1 is preformed and rigid. The disaccharide Xyl-Glc- in the acceptor ligand is shown as spheres, and the EGF motif is shown as semi-transparent gray surface. The disulfide bond between C356 and C385, the side chains of H262, W265, G325, W359, and W358 around the active site pocket are shown as sticks. The square box marks a 4-residue loop region (G194–T197) with notable conformational change, but this region is far away from the donor and the acceptor ligands.




**Supplementary Figure 5. Crystal contacts in the XXYLT1/Xyl-Glc-EGF binary complex around EGF moiety.** Closest contacts are shown in the zoomed views of contact areas A and B, where the orange structure is superimposed XXYLT1/$Mn^{2+}$ structure determined in the absence of the acceptor Xyl-Glc-EGF.

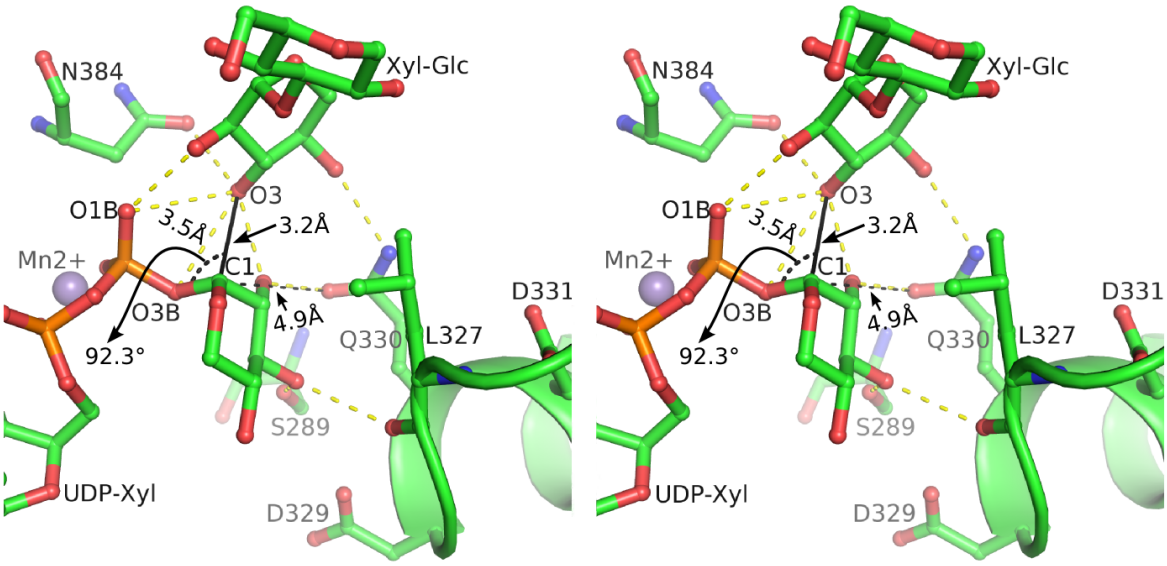



**Supplementary Figure 6. Site-directed mutagenesis of XXYLT1 and *in vitro* activity assays of the mutant enzymes.** (**a**) The expression of Myc/$His_6$-tagged wild type and mutant XXYLT1 proteins in transiently transfected HEK293T cells was confirmed by the Western blot analysis of culture media (top panel) and cell lysates (middle panel) with anti-Myc monoclonal antibody. The expression of β-actin was used as loading control (bottom). (**b**) The xylosyltransferase activities of wild type XXYLT1 against hFA9 Xyl-Glc-EGF (shown as No Fucose), or hFA9 Xyl-Glc-EGF additionally modified with the *O*-fucose monosaccharide (Fucose), or the *O*-fucose disaccharide (GlcNAc-Fucose) at S61 of the protein. (**c**) Buffer-dependent xylosyltransferase activities of wild type XXYLT1. *In vitro* activity was tested under Bis-tris buffer and HEPES buffer. The data in (**b** and **c**) were from three independent assays. The bars indicate mean ± S.E.M.

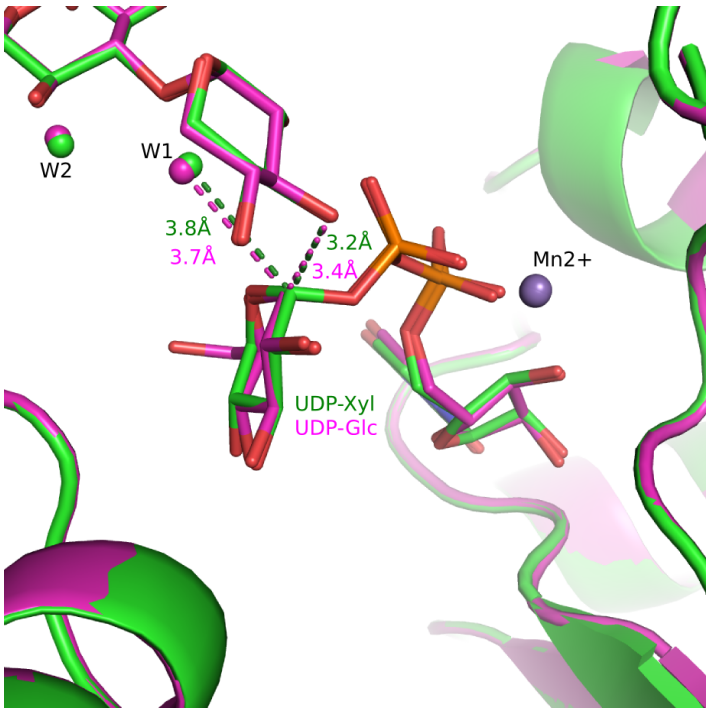



**Supplementary Figure 7. The gradual appearance of an alternative UDP conformation as a proxy for the reaction state.** (**a**) The active site structure of Product complex II (shown as blue sticks). The yellow dashed lines indicate H-bonds. The black dashed lines showed shorter-than 3.2 Å bump distances between the enzyme and the transferred xylose, demonstrating a highly stressed local environment. (**b**) In Product complex III (shown as salmon sticks), the unbiased *Fo-Fc* electron density map (3σ) revealed two alternating conformations of UDP, likely a result of relieving the steric

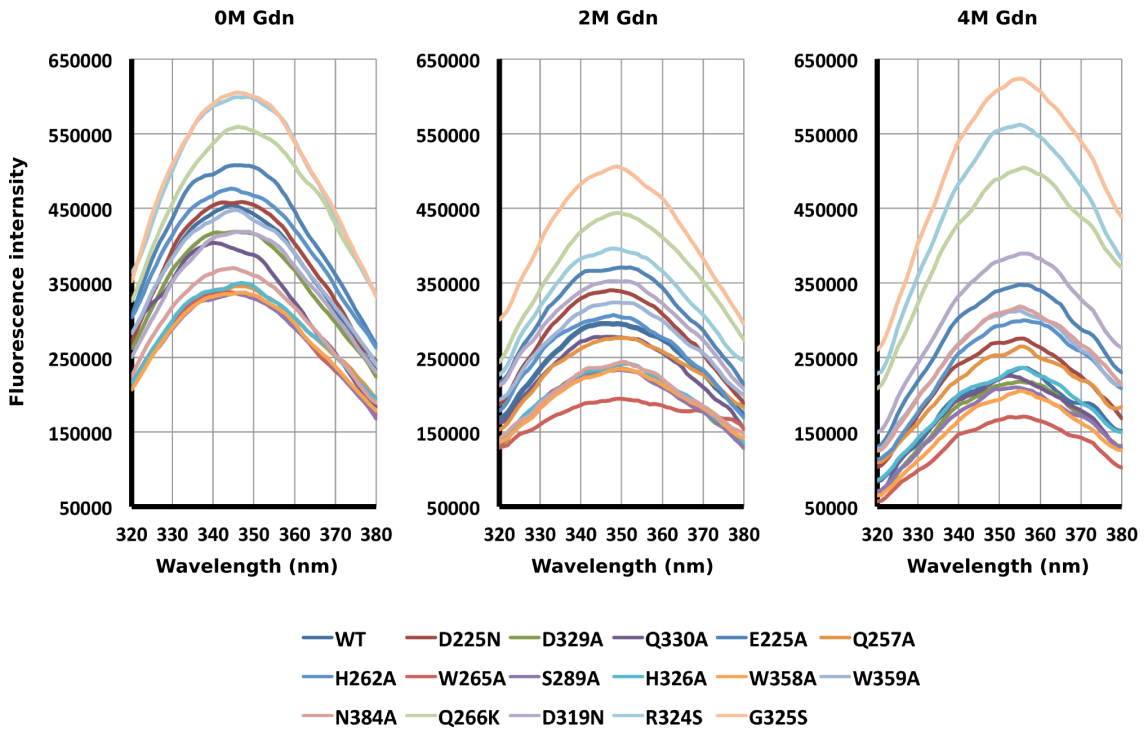
stress from the terminal xylose of the final product Xyl-Xyl-Glc. The superposed Product complex II are shown as blue sticks. (**c**) Superposition of the two UDP conformations of Product complex III (in stick view) into the unbiased *Fo-Fc* difference densities of UDP-Xyl Michaelis complex (first panel), Product complex I (second panel), Product complex II (third panel), Product complex III (forth panel), UDP-Glc Michaelis complex (fifth panel), and UDP complex (sixth panel), respectively. The density marked by a circle gradually developed as the transferring reaction proceeded. The encircled region was occupied by the α-phosphate in the alternative UDP conformation, which contained full density in the completion Product complex III, partial density in Product complex II, less density in Product complex I, and largely empty in UDP-Xyl Michaelis complex, UDP-Glc Michaelis complex, and UDP complex. (**d)** Superposition of the UDP moieties (stick view) in Product complex I (cyan), Product complex II (blue), Product complex III (salmon), and UDP complex (orange). As the transfer of xylose from UDP to acceptor disaccharide proceeded, the pyrophosphate of the UDP moiety gradually moved away from xylose as indicated by the three black arrows. **(e)** The pyrophosphate positions in UDP-Xyl Michaelis complex (green) and UDP-Glc Michaelis complex (magenta) are closer to the pyrophosphate position of Product Complex I (cyan) than to that of Product Complex II (blue) and UDP complex (orange), in consistent with their respective proposed states. Note that the 2.8 Å distance (yellow dashed line) between UDP pyrophosphate and acceptor O3 was significantly shorter than the 3.5 Å distance in the intact UDP-Xyl or 3.6 Å distance in UDP-Glc, suggesting that the pyrophosphates of the intact donor ligands was in a “pressed-down” conformation. The stressed UDP conformation may facilitate the cleavage of the donor sugar from the pyrophosphate during the transfer reaction and subsequent departure of the product.



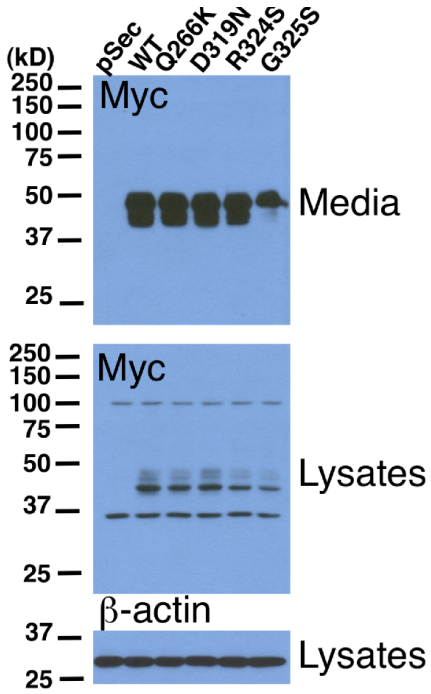

**Supplementary Figure 8. Stereo view of the active site structure and transfer geometry of the UDP-Xyl Michaelis ternary complex.** Yellow dashed lines mark H-bonds. The distances from donor Xyl anomeric carbon to either acceptor Xyl O3 hydroxyl (3.2 Å) or Q330 amide oxygen (4.9 Å) are marked by black solid and black dashed lines, respectively. $Mn^{2+}$ is shown as a purple sphere.




**Supplementary Figure 9. Waters are less likely to participate in the transfer reaction.** In the UDP-Xyl Michaelis complex (Green) and UDP-Glc Michaelis (Magenta), only two waters closest to their anomeric carbons are shown.




**Supplementary Figure 10. Effects of XXYLT1 mutations on XXYLT1 structure were monitored by intrinsic tryptophan fluorescence.** Tryptophan residues were excited at 280 nm, and fluorescence was recorded over the range of 320–380 nm. One representative data set from two independent assays was shown.




**Supplementary Figure 11. The expression of Myc/$His_6$-tagged wild type and selected cancer-related mutants of XXYLT1 in transiently transfected HEK293T cells**. Protein expression was confirmed by the Western blot analysis of culture media (top panel) and cell lysates (middle panel) with anti-Myc monoclonal antibody. The expression of β-actin was used as loading control (bottom).

## Supplementary Table 1. Data collection and structure refinement statistics of the seven XXYLT1-related structures.

| Complex | Xxylt1:$Mn^{2+}$ | Xxylt1:Xyl-Glc -EGF | UDP-Xyl Michaelis complex (10 min) | Product complex I (20 min) | Product complex II (20 min) | Product complex III (60 min) | UDP-Glc Michaelis complex (60 min) | UDP complex (60 min) | $Sm^{3+}$ derivative |
|---|---|---|---|---|---|---|---|---|---|
| Ligand | $Mn^{2+}$ | Xyl-Glc-EGF | Xyl-Glc-EGF, UDP-Xyl, $Mn^{2+}$ | Xyl-Xyl-Glc-EGF, UDP, $Mn^{2+}$ | Xyl-Xyl-Glc-EGF, UDP, $Mn^{2+}$ | Xyl-Xyl-Glc-EGF, UDP, $Mn^{2+}$ | Xyl-Glc-EGF, UDP-Glc, $Mn^{2+}$ | Xyl-Glc-EGF, UDP, $Mn^{2+}$ | $Sm^{3+}$ |
| PDB ID | 4WLM | 4WM0 | 4WNH | 4WMI | 4WMK | 4WN2 | 4WMA | 4WMB | - |
| **Data collection** | | | | | | | | | |
| Space group | $P3_121$ | $P3$ | $P3$ | $P3$ | $P3$ | $P3$ | $P3$ | $P3$ | $P3_121$ |
| Cell dimensions: | | | | | | | | | |
| *a, b, c* (Å) | 89.31, 89.31,154.30 | 89.54, 89.54, 42.98 | 89.48, 89.48, 42.90 | 89.06, 89.06, 42.74 | 88.91, 88.91, 42.68 | 89.51,89.51, 43.05 | 90.03,90.03, 43.14 | 89.47,89.47, 42.93 | 91.06, 91.06,153.99 |
| *α, β, γ* (°) | 90, 90, 120 | 90,90,120 | 90,90,120 | 90,90,120 | 90,90,120 | 90,90,120 | 90,90,120 | 90,90,120 | 90,90,120 |
| Resolution (Å) | 40-3.00(3.11-3.00) | 50-2.37 (2.45-2.37) | 50-1.95 (2.02-1.95) | 40-1.87 (1.94-1.87) | 40-2.08 (2.15-2.08) | 50-1.95 (2.02-1.95) | 50-1.62 (1.68-1.62) | 50-2.05 (2.12-2.05) | 50-3.58 (3.71-3.58) |
| Rmerge (%) | 12.2 (58.9) | 6.1 (19.6) | 7.0 (65.1) | 8.4 (58.0) | 7.0 (41.2) | 4.9 (40.2) | 4.5 (59.2) | 5.3 (42.9) | 23.5 (57.9) |
| *I/σ(I)* | 15.6 (2.8) | 22.6 (6.5) | 19.5 (2.3) | 19.2 (2.0) | 14.0 (2.3) | 16.3 (2.1) | 28.2 (2.2) | 17.1 (2.5) | 6.9 (2.0) |
| Completeness (%) | 99.8 (99.9) | 99.6 (96.0) | 100 (100) | 99.6 (96.4) | 97.8 (84.1) | 99.4 (93.8) | 99.9 (98.8) | 99.9 (99.0) | 99.5 (94.9) |
| Redundancy | 5.9 (6.1) | 4.8 (3.9) | 4.8 (3.7) | 8.9 (4.6) | 4.6 (3.5) | 4.7 (3.1) | 9.6 (6.6) | 4.8 (4.0) | 8.7 (6.6) |
| **Refinement** | | | | | | | | | |
| Resolution (Å) | 40-3.00 | 50-2.37 | 50-1.95 | 40-1.87 | 40-2.08 | 50-1.95 | 50-1.62 | 50-2.05 | |
| No. reflections | 14806 | 15616 | 28026 | 31390 | 22674 | 28075 | 49949 | 24139 | |
| $R_{work}/R_{free}$ (%) | 23.0/28.7 | 18.8/23.3 | 19.3/22.5 | 19.5/22.6 | 19.1/24.6 | 19.5/23.8 | 21.9/24.3 | 20.1/23.5 | |
| No. atoms | | | | | | | | | |
| Protein | 4747 | 2440 | 2435 | 2440 | 2440 | 2440 | 2440 | 2440 | |
| EGF | - | 263 | 263 | 263 | 263 | 263 | 263 | 263 | |
| Nucleotide(sugar)/$Mn^{2+}$/ waters | - /2/ - | 20/ - / 76 | 54/1/80 | 54/1/120 | 54/1/86 | 54/1/92 | 56/1/135 | 45/1/71 | |
| *B*-factors | | | | | | | | | |
| Protein | 56 | 34 | 32 | 29 | 31 | 34 | 30 | 37 | |
| EGF | - | 54 | 51 | 43 | 48 | 57 | 49 | 56 | |
| Nucleotide(sugar)/$Mn^{2+}$/ Waters | - /43/ - | 28/ - /35 | 29/24/34 | 32/23/35 | 34/23/33 | 33/26/35 | 25/21/34 | 32/27/37 | |
| R.m.s deviations | | | | | | | | | |
| Bond lengths (Å) | 0.006 | 0.007 | 0.008 | 0.009 | 0.009 | 0.011 | 0.009 | 0.008 | |
| Bond angles (°) | 0.917 | 1.196 | 1.278 | 1.399 | 1.375 | 1.527 | 1.370 | 1.260 | |

**Supplementary Table 2. Intrinsic Tryptophan fluorescence features of XXYLT1 and its mutants under 0 M, 2 M and 4 M guanidinium chloride (Gdn).**

| Proteins | 0M Gdn | | 2M Gdn | | 4M Gdn | |
|---|---|---|---|---|---|---|
| | 340/350 nm | λ max | 340/350 nm | λ max | 340/350 nm | λ max |
| WT | 1.001 | 344.8 | 0.975 | 347 | 0.889 | 356.0 |
| D225N | 1.001 | 344.8 | 0.972 | 347 | 0.869 | 354.5 |
| D329A | 1.000 | 345.0 | 0.961 | 348 | 0.866 | 355.5 |
| Q330A | 1.034 | 341.5 | 0.980 | 348 | 0.865 | 354.0 |
| E225A | 0.994 | 344.0 | 0.976 | 347 | 0.874 | 355.5 |
| Q257A | 0.996 | 343.5 | 0.946 | 349 | 0.856 | 354.5 |
| H262A | 1.001 | 344.0 | 0.972 | 348 | 0.859 | 355.0 |
| W265A | 1.023 | 342.5 | 0.958 | 350 | 0.864 | 354.5 |
| S289A | 1.007 | 344.5 | 0.968 | 347 | 0.836 | 354.5 |
| H326A | 0.995 | 343.5 | 0.944 | 350 | 0.858 | 355.0 |
| W358A | 1.003 | 344.5 | 0.937 | 350 | 0.850 | 355.5 |
| W359A | 1.001 | 345.3 | 0.960 | 348 | 0.874 | 354.5 |
| N384A | 1.004 | 344.5 | 0.943 | 350 | 0.859 | 355.5 |
| Q266K | 0.975 | 347.0 | 0.944 | 349 | 0.884 | 355.5 |
| D319N | 0.985 | 346.5 | 0.963 | 349 | 0.869 | 356.0 |
| R324S | 0.990 | 346.5 | 0.969 | 348 | 0.868 | 355.5 |
| G325S | 0.987 | 346.0 | 0.957 | 349 | 0.861 | 355.5 |
| AVG | 1.000 | 344.6 | 0.960 | 348 | 0.865 | 355.1 |
| SD | 0.014 | 1.4 | 0.013 | 1.1 | 0.012 | 0.6 |

**Supplementary Table 3. Compilation of XXYLT1 mutations identified in some cancers.** Gene alteration data are obtained from cancer genomics site**:** cBioPortal (http://www.cbioportal.org)**.**

| hXXYLT1 Mutations | Corresponding residue of mXXYLT1 | Cancer types |
|---|---|---|
| A121E | A120 | Lung Adenocarcinoma |
| S208L | S207 | Stomach Adenocarcinoma |
| R236L | R235 | Lung Adenocarcinoma |
| F242L | F241 | Breast Invasive Carcinoma |
| D243H | D242 | Lung Squamous Cell Carcinoma |
| G252S | G251 | Liver Hepatocellular Carcinoma |
| Q267K | Q266 | Breast Invasive Carcinoma |
| R276W | R275 | Stomach Adenocarcinoma |
| G279V | D278 | Small Cell Lung Cancer |
| P280S | P279 | Skin Cutaneous Melanoma |
| A299S | A298 | Head and Neck Squamous Cell Carcinoma |
| R301H | R300 | Uterine Corpus Endometrial Carcinoma |
| R308C | H307 | Breast Invasive Carcinoma |
| D320N | D319 | Multiple Myeloma |
| Y322C | Y321 | Stomach Adenocarcinoma |
| R325S | R324 | Liver Hepatocellular Carcinoma |
| G326S | G325 | Stomach Adenocarcinoma |
| G329R | G328 | Stomach Adenocarcinoma |
| H378P | H377 | Lung Adenocarcinoma |
| V379I | V378 | Bladder Urothelial Carcinoma |
| Y382C | Y381 | Stomach Adenocarcinoma |
| D393N | D392 | Skin Cutaneous Melanoma |

## Supplementary Table 4. A list of primers used for site-directed mutagenesis

| | | |
|---|---|---|
| D225N | Forward: | 5'-TCCAGCTGAACCTTGACCTGAAGTATAAGACCAAC-3' |
| | Reverse: | 5'-GTCAAGGTTCAGCTGGATGATCCTCGGGATC-3' |
| E255A | Forward: | 5'-AGCCAGAGCGATGCAGCCTGTGTACAGGCACAC-3' |
| | Reverse: | 5'-GCTGCATCGCTCTGGCTATGCCGATAACAGCGC-3' |
| Q257A | Forward: | 5'-GAGAGATGGCGCCTGTGTACAGGCACACGTTC-3' |
| | Reverse: | 5'-ACACAGGCGCCATCTCTCTGGCTATGCCGATAAC-3' |
| H262A | Forward: | 5'-TGTACAGGGCCACGTTCTGGCAGTTCCGCCAT-3' |
| | Reverse: | 5'-AGAACGTGGCCCTGTACACAGGCTGCATCTCTCT-3' |
| W265A | Forward: | 5'-ACACGTTCGCGCAGTTCCGCCATGAGAACCCC-3' |
| | Reverse: | 5'-GGAACTGCGCGAACGTGTGCCTGTACACAGGCT-3' |
| Q266K | Forward: | 5'-CGTTCTGGAAGTTCCGCCATGAGAACCCC-3' |
| | Reverse: | 5'-GCGGAACTTCCAGAACGTGTGCCTGTACAC-3' |
| N288A | Forward: | 5'-CTGGCTTCGCCAGTGGAGTGATGTTGCTGAACCTG-3' |
| | Reverse: | 5'-CTCCACTGGCGAAGCCAGGGAGTCCTTCAGGCGG-3' |
| S289A | Forward: | 5'-GCTTCAACGCTGGAGTGATGTTGCTGAACCTGGAGG-3' |
| | Reverse: | 5'-TCACTCCAGCGTTGAAGCCAGGGAGTCCTTCAG-3' |
| D319N | Forward: | 5'-CAGCTTGCTAACAAGTACCACTTCCGGGG-3' |
| | Reverse: | 5'-GGTACTTGTTAGCAAGCTGCTGTACCCACGA-3' |
| R324S | Forward: | 5'-ACCACTTCTCGGGCCACCTGGGGGACCA-3' |
| | Reverse: | 5'-GGTGGCCCGAGAAGTGGTACTTGTCAGCAAGC-3' |
| G325S | Forward: | 5'-ACTTCCGGAGCCACCTGGGGGACCAG-3' |
| | Reverse: | 5'-CAGGTGGCTCCGGAAGTGGTACTTGTCAGC-3' |
| H326A | Forward: | 5'-TCCGGGGCGCCCTGGGGGACCAGGACTTCTT-3' |
| | Reverse: | 5'-CCCAGGGCGCCCCGGAAGTGGTACTTGTCAGCAAGCTGC-3' |
| D329A | Forward: | 5'-CCTGGGGGCCCAGGACTTCTTCACCATGATTGGC-3' |
| | Reverse: | 5'-AGTCCTGGGCCCCCAGGTGGCCCCGGAA-3' |
| Q330A | Forward: | 5'-TGGGGGACGCGGACTTCTTCACCATGATTGGCATG-3' |
| | Reverse: | 5'-AGAAGTCCGCGTCCCCCAGGTGGCCCC-3' |
| D331A | Forward: | 5'-GGACCAGGCCTTCTTCACCATGATTGGCATG-3' |
| | Reverse: | 5'-TGAAGAAGGCCTGGTCCCCCAGGTGG-3' |
| C356S | Forward: | 5'-GCAGCTGTCCACCTGGTGGAGGGACCATGGCTA-3' |
| | Reverse: | 5'-ACCAGGTGGACAGCTGCCGGTTCCAGGTGCAGT-3' |
| W358A | Forward: | 5'-TGTGTACCGCGTGGAGGGACCATGGCTACAGCGATGT-3' |
| | Reverse: | 5'-CCCTCCACGCGGTACACAGCTGCCGGTTCCAG-3' |
| W359A | Forward: | 5'-GTACCTGGGCGAGGGACCATGGCTACAGCGATGTCTTC-3' |
| | Reverse: | 5'-GGTCCCTCGCCCAGGTACACAGCTGCCGGTTCCA-3' |
| N384A | Forward: | 5'-ACCATGGGGCCTGCAACACACCCATCCCAGAG-3' |
| | Reverse: | 5'-TGTTGCAGGCCCCATGGTAGATCTTGACGTGGCCCT-3' |
| C385S | Forward: | 5'-TGGGAACTCCAACACACCCATCCCAGAGGAC-3' |
| | Reverse: | 5'-GTGTGTTGGAGTTCCCATGGTAGATCTTGACGTGG-3' |
| *D329/331A | Forward: | 5'-CCTGGGGGCCCAGGCCTTCTTCACCATGATTGGC-3' |
| | Reverse: | 5'-AGGCCTGGGCCCCCAGGTGGCCCCGGAA-3' |

*Double mutant. The second PCR for incorporation of D329A mutation was performed using the plasmid encoding the D331A mutant as template.

**Supplementary Movie 1. A movie morphing the isolated hFA9 EGF (PDB 1EDM) to the same protein in complex with XXYLT1.**
We used the 'Morph' function in Chimera to generate the movie coordinates and displayed the movie in PyMOL. The Xyl-Glc disaccharide in the starting structure was modeled on the isolated hFA9 EGF structure based on a glucose modified EGF structure (PDB ID: 4XL1). The disaccharide was modeled away from EGF to prevent the steric clashes between the apical Xylose and EGF.

**Supplementary Movie 2. Morphing (generated using Chimera) of the electron densities of donor and acceptor substrates of the four trapped ternary complexes.**
They are UDP-Xyl Michaelis complex, Product complex I, II and III as labeled in this movie. We used the unbiased *Fo-Fc* difference densities calculated before ligands were modeled. The electron density maps were aligned and displayed at 3.0σ in the movie. Three blue arrows highlight structural features that are consistent with an ongoing transfer reaction, and the lower left arrow marks the shift of pyrophosphate as described in the main text.