

Supporting Information

Chaikuad et al. 10.1073/pnas.1113921108

SI Materials and Methods

Expression and Purification of hGYG1. hGYG1 was cultured in 3 L of Terrific Broth at 37 °C, and induced with 0.1 mM IPTG overnight at 18 °C. Cell pellets were harvested, homogenized in lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol, 10 mM imidazole) and insoluble material was removed by centrifugation. The supernatant was purified by affinity (Ni-Sepharose; GE Healthcare) and size-exclusion (Superdex 75; GE Healthcare) chromatography. Purified protein was treated with His-tagged TEV protease overnight at 4 °C, and then passed over 0.5 mL of Ni-Sepharose resin. The purified hGYG1 protein was concentrated to 10 mg/mL and stored in storage buffer (10 mM HEPES pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, 0.5 mM TCEP) at –80 °C.

In Vitro Glucosylation Assay. Five μ M purified hGYG1 enzyme was incubated in 25 mM HEPES pH 7.5, 1 mM MnCl₂ and 1 mM UDP-glucose for either 0 or 60 minutes at 37 °C, before the reaction was terminated with the addition of Na-EDTA at a final concentration of 20 mM. Protein intact mass was then determined by using an MSD-TOF electrospray ionization orthogonal time-of-flight mass spectrometer and analyzed by TOF Protein Confirmation Software (Agilent).

Crystallization and Data Collection. Crystals were grown by sitting drop vapour diffusion at 20 °C. The crystallization conditions for various hGYG1 structures are summarized in Table S1. To crystallize hGYG1-liganded complexes, protein (10 mg/mL)

was preincubated with 3 mM of the corresponding ligand. Crystals were cryoprotected with reservoir solution supplemented with 25% (v/v) glycerol or ethylene glycol, and flash-cooled in liquid nitrogen. Diffraction data were collected on the Diamond Light Source beamlines or in-house Rigaku FR-E-Superbright, and processed with XDS (1), or MOSFLM and SCALA from the CCP4 Suite (2).

Structure Determination. Structures were solved by molecular replacement with PHASER (3) using the rat rGYG1 coordinates as a search model. Density modification and NCS averaging were performed with PARROT (4), and the improved phases were used for automated model building with ARP/wARP (5). Iterative cycles of manual model building using COOT (6) alternated with refinement in REFMAC5 (7) were performed. Optimal TLS groups calculated by the TLSMD server (8) were used in the last refinement step. Data collection, refinement statistics, and Protein Data Bank (PDB) ID codes are summarized in Table S1. In the structures of hGYG1_{WT-mix} complexed with Mn²⁺ and UDP, electron density provided clear evidence of a maltosaccharide chain locked in the active site of each hGYG1 protomer. However, this density feature can vary in occupancy and shape, possibly representing an average of different maltosaccharides in the crystal that reflects the chain length heterogeneity in the hGYG1_{WT-mix} protein. However, after characterizing a number of crystals we obtained two structures that allow full tracing of Tyr195 (O)-linked maltotetraose and maltohexaose chains.

1. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* 66:133–144.
2. CCP4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50:760–763.
3. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* 61:458–464.
4. Zhang KY, Cowtan K, Main P (1997) Combining constraints for electron-density modification. *Methods Enzymol* 277:53–64.
5. Perrakis A, Morris R, Lamzin VS (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* 6:458–463.
6. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.
7. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53:240–255.
8. Painter J, Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* 62:439–450.
9. Paul D, et al. (2010) Glycal formation in crystals of uridine phosphorylase. *Biochemistry* 49:3499–3509.
10. Smar M, Short SA, Wolfenden R (1991) Lyase activity of nucleoside 2-deoxyribosyltransferase: Transient generation of ribal and its use in the synthesis of 2'-deoxynucleosides. *Biochemistry* 30:7908–7912.

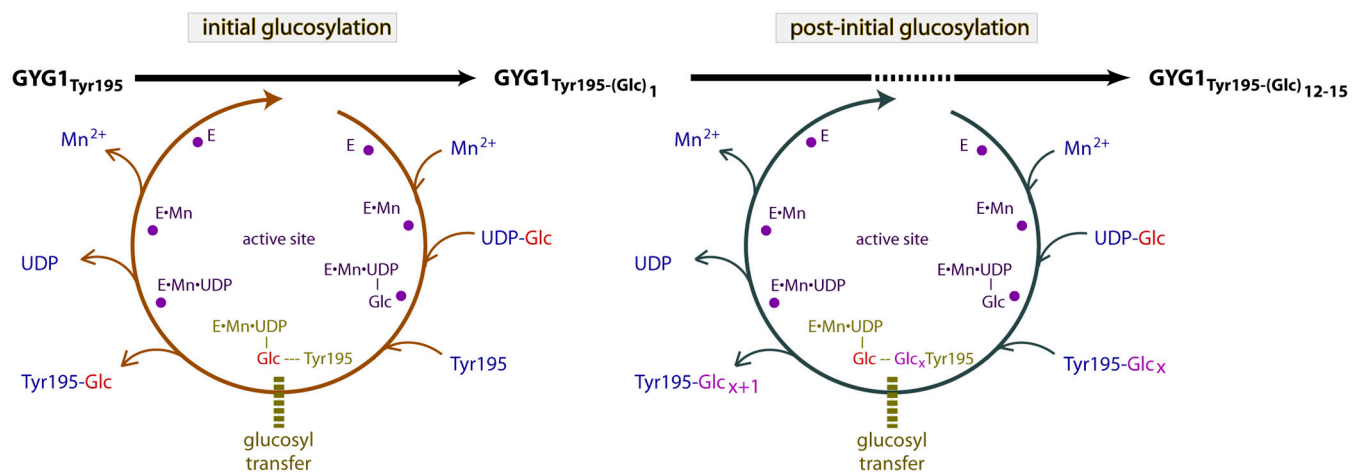


Fig. S1. Schematic diagram of glucosyl transfer reaction catalyzed by glycogenin. Glucosyl transfer catalyzed by glycogenin is believed to follow an ordered bi-bi mechanism, as proposed for other GTs, with the sequential binding of Mn^{2+} , donor (UDP-Glc) and acceptor (Tyr195-Glc_x). There are two types of glucosyl transfer for GYG1. At the initial glucosylation step, it transfers the first glucose unit onto its own acceptor, Tyr195, forming a C1-O-tyrosyl linkage. In the subsequent, sequential glucosyl transfer, the terminal glucose acts as an acceptor forming an α -1,4-glucosyl linkage to an incoming glucose (from a new UDP-Glc donor). The postinitial glucosylation step is repeated, allowing a nascent maltosaccharide chain to elongate until it has reached 12–15 glucose units. The enzyme then exits the cycle and the attached maltosaccharide chain serves as a substrate for glycogen synthase and branching enzyme.

Protein sources	ligands	crystal properties		biological features				structure names
		crystal lattices	no. protein in AU	enzyme state	lid	donor subsite	acceptor subsite	
hGYG1 _{WT-mix}	Mn ²⁺ UDP	triclinic	2	active	closed	4x-glucose chain		hGYG1 _{WT-mix} • Mn ²⁺ • UDP • Glc ₄
	Mn ²⁺ UDP	triclinic	2	active	closed	6x-glucose chain		hGYG1 _{WT-mix} • Mn ²⁺ • UDP • Glc ₆
	Mn ²⁺ /	orthorhombic	1	ground	open	/	/	hGYG1 _{WT-mix} • Mn ²⁺
hGYG1 _{WT-0}	Mn ²⁺ UDPG	triclinic	2	active	closed	glucose	glucose & β-water	hGYG1 _{WT-0} • Mn ²⁺ • UDPG
	Mn ²⁺ UDP	triclinic	2	active	closed	/	/	hGYG1 _{WT-0} • Mn ²⁺ • UDP (triclinic)
	Mn ²⁺ UDP	monoclinic	2	active	partially closed	/	/	hGYG1 _{WT-0} • Mn ²⁺ • UDP (monoclinic)
hGYG1 _{Y195F}	/ /	orthorhombic	1	ground	open	/	/	hGYG1 _{Y195F} • apo
	Mn ²⁺ UDP	orthorhombic	1	ground	open	/	/	hGYG1 _{Y195F} • Mn ²⁺ • UDP
	Mn ²⁺ UDPG	monoclinic	2	active	closed	glucosyl species	α-water	hGYG1 _{Y195F} • Mn ²⁺ • UDPG
hGYG1 _{T83M}	Mn ²⁺ UDPG	orthorhombic	1	ground	open	glucose	/	hGYG1 _{T83M} • Mn ²⁺ • UDPG
	Mn ²⁺ UDP	orthorhombic	1	ground	open	/	/	hGYG1 _{T83M} • Mn ²⁺ • UDP

Fig. S2. Summary of hGYG1 structures determined in this study. Using different sources of recombinant proteins in combination with ligands, hGYG1 demonstrates the ability to crystallize in various chemical spaces, resulting in at least three crystal forms that correlate with the major structural features of the enzyme, namely the lid conformation (open, partially closed, closed), enzyme state (ground, active) and occupants at the donor and acceptor subsites. AU, asymmetric unit.

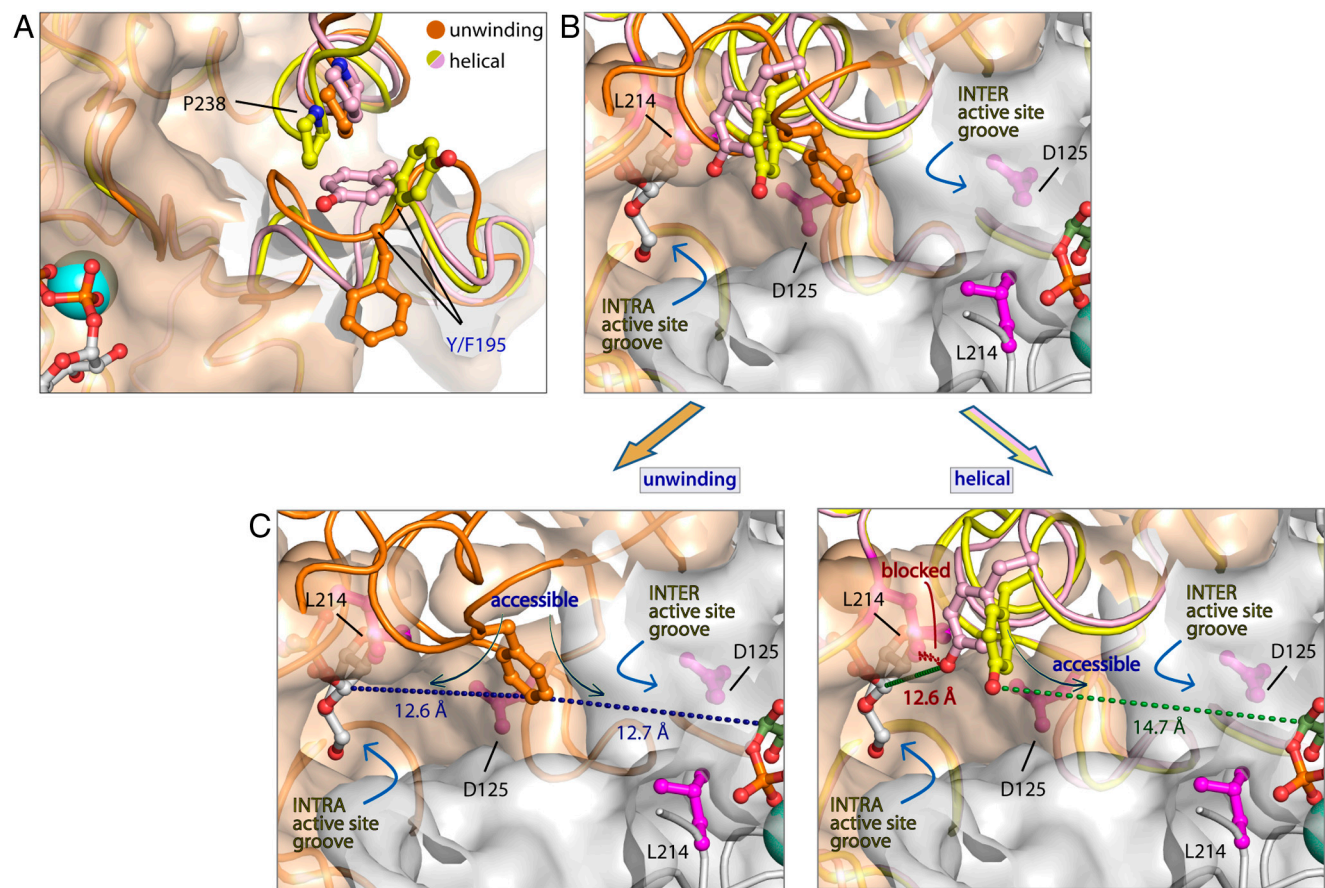


Fig. 53. Conformational plasticity of acceptor arm and Tyr195. (A) While disordered in the ground state, the acceptor arm is observed to adopt either an unwinding coil (orange, in $\text{hYG}1_{\text{Y195F}} \cdot \text{Mn}^{2+} \cdot \text{UDPG}$ structure) or helical conformation (yellow and pink, in $\text{hYG}1_{\text{WT-0}} \cdot \text{Mn}^{2+} \cdot \text{UDPG}$ and $\text{hYG}1_{\text{WT-mix}} \cdot \text{Mn}^{2+} \cdot \text{UDP} \cdot \text{Glc}_4$ structures, respectively) in the active state. This phenomenon is mediated by different arrangements of the C loop and its residues, mainly Pro238, that are juxtaposed to the acceptor arm, and hence both regions typically display conformations complementary to each other. (B) Close-up of the Tyr195 acceptor residue (yellow and pink) or the counterpart Phe195 in the $\text{hYG}1_{\text{Y195F}}$ mutant (orange). Comparison of the acceptor residue positions between the unwinding and helical conformations reveals a similar location within the dimeric interface. (C) Bottom left: When the acceptor arm unwinds, the acceptor residue is equidistant (approximately 13 Å) to both active sites in the dimer. Bottom right: When the acceptor arm forms a helical structure, Tyr195 is observed in two conformations. In an "inward" conformation (pink), Tyr195 lies closest to its own active site (linear distance of 12.6 Å), but a linear approach to this active site is sterically hindered by Leu214 (red squiggly line). When pointing "outward" (yellow), Tyr195 is closest to the active site in the opposite subunit (14.7 Å), and a linear approach is not sterically hindered.

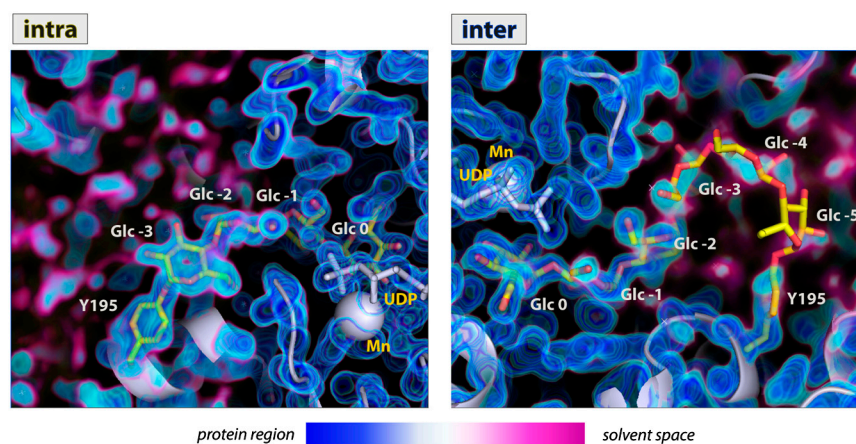


Fig. 54. Illustration of the protein boundary and solvent exterior in proximity to acceptor arm and active site for intra- and intersubunit threading modes. Inspection of the attached maltosaccharide chains in the $\text{hYG}1$ product complexes ($\text{hYG}1_{\text{WT-mix}} \cdot \text{Mn}^{2+} \cdot \text{UDP} \cdot \text{Glc}_4$ or $\text{hYG}1_{\text{WT-mix}} \cdot \text{Mn}^{2+} \cdot \text{UDP} \cdot \text{Glc}_6$ structures) highlights differences in their paths with respect to the protein boundary and solvent exterior. A short maltosaccharide chain (e.g., Glc_3) can traverse a short path into the active site in an intrasubunit threading mode, with all glucose units packed and buried along the grooves at the protein surface (left). However, a longer maltosaccharide chain (e.g., Glc_5 or longer) can adopt an intersubunit threading mode that exposes most of its glucose units to the solvent space, leaving only the terminal two glucose units (Glc-1 and Glc-2) within the protein boundary (right), close to the donor glucose (at Glc 0). The protein and solvent regions are calculated using PyMol based on electron density map.

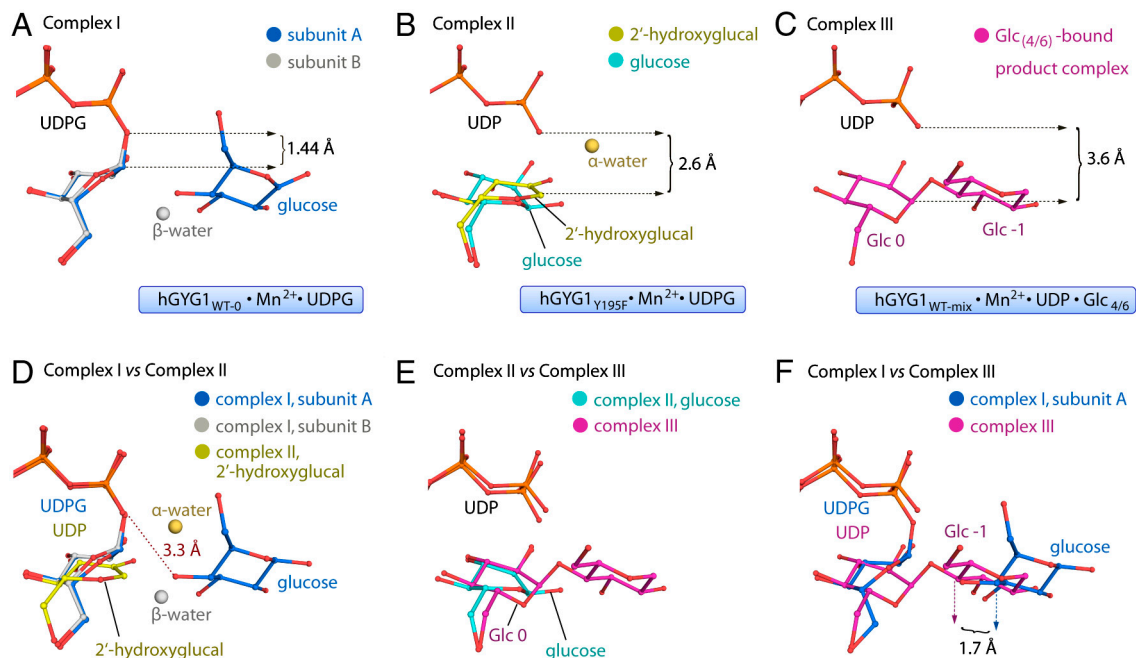


Fig. S5. Atomic coordination of glucose moieties in the donor and acceptor subsites in different hGYG1 complexes. When we cocrystallized hGYG1 with Mn^{2+} and UDPG, two distinct complex structures bound with different ligands at the donor and acceptor subsites were resolved. (A) In complex I ($hGYG1_{WT-0} \cdot Mn^{2+} \cdot UDPG$), the glucose moiety from UDPG is bound in the donor subsite, tucked below the UDP-pyrophosphate. The linkage between the β -phosphate O3 and glucose C'1 remains intact. The acceptor subsite contains, in one subunit a water molecule (" β -water," gray) at the β -face (i.e., opposite to the C'1-O3 glucose-pyrophosphate scissile bond of UDPG) and in another subunit a glucose molecule (blue), presumably a hydrolysis product of UDPG. (B) In complex II ($hGYG1_{Y195F} \cdot Mn^{2+} \cdot UDPG$), the electron density at the donor subsite reveals a bound ligand clearly separated from the UDP (i.e., no covalent linkage). While its exact identity remains unclear, we have modeled two possible glucosyl species that are compatible with the electron density: 2'-hydroxyglucal (yellow) and glucose (cyan). Glucose would likely derive from UDPG hydrolysis, while 2'-hydroxyglucal might be a product from an elimination reaction of the unstable oxocarbenium intermediate. Enzyme-catalyzed glycal formation has been observed previously in the characterization of some glycosyltransferases and glycoside hydrolases (9, 10). The position of our putative 2'-hydroxyglucal species is supported from a structure of $hGYG1 \cdot Mn^{2+} \cdot UDP \cdot 1'$ -deoxy-glucose in this study (PDB ID code 3U2X), which showed that 1'-deoxy-glucose, with similar chemical compositions to 2'-hydroxyglucal, is bound in the same position at the donor subsite. The acceptor subsite in complex II contains a water molecule (" α -water") at the α -face (i.e., same side of the UDPG scissile bond). (C) The maltosaccharide-bound complexes (complex III, $hGYG1_{WT-mix} \cdot Mn^{2+} \cdot UDP \cdot Glc_4$ or $hGYG1_{WT-mix} \cdot Mn^{2+} \cdot UDP \cdot Glc_6$) map positions of donor and acceptor glucose after glucosyl transfer. Comparison of these three complexes suggests possible atomic coordination motions of glucose moieties during glucosyl transfer. (D) The yellow glucosyl species in complex II lies in the same vertical plane as the glucose of UDPG in complex I (gray and blue) but displaces downward with its anomeric C'1 atom 2.6 Å away from the UDP β -phosphate O3, indicative of a possible donor glucose position in the transition state after its departure from UDP. In this position, the donor glucose C'1 atom is equi-planar to the O'4 atom of the acceptor glucose in complex I, which can be activated by the leaving UDP β -phosphate (red dotted line) to attack in an α -retaining manner. (E) The cyan glucosyl species in complex II aligns well with Glc 0 unit of complex III (magenta), suggesting it may represent the position of the glucose after being transferred. (F) Comparison of the glucose in the acceptor subsite of complex I (blue) and complex III (Glc -1, magenta) reveal a displacement of 1.7 Å. This demonstrates that the attacking glucose acceptor at the initial binding position (blue) has to move inward (i.e., toward the donor glucose) after forming an α -1,4 linkage with the donor glucose.

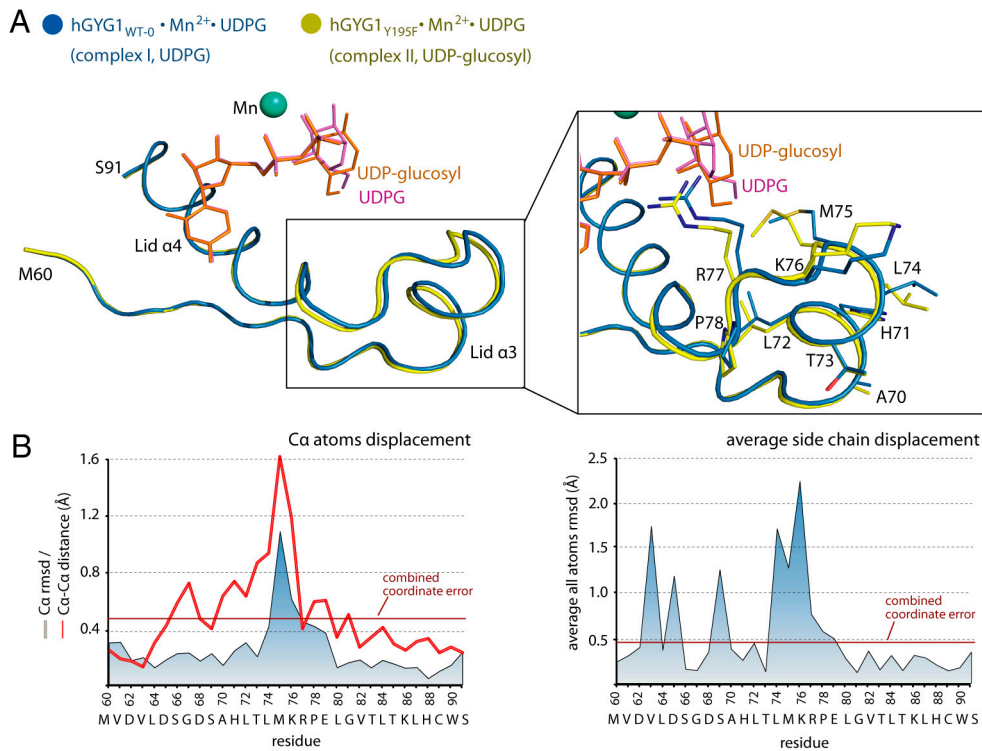
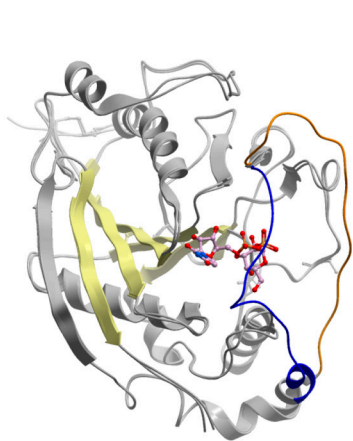
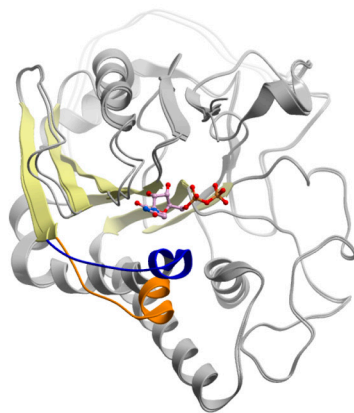


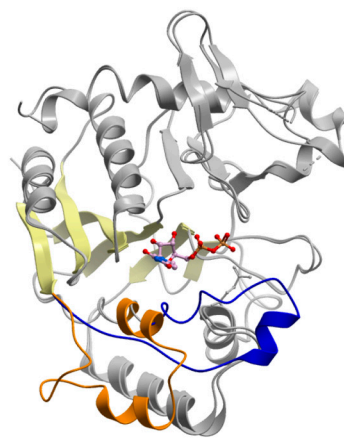
Fig. S6. Adjustment of Met75 and residues in helix α_3 of the closed lid. (A) Comparison between the closed lid in hGYG1_{WT-0} · Mn²⁺ · UDPG (blue) and hGYG1_{Y195F} · Mn²⁺ · UDPG (yellow) structures demonstrates small alterations of the helix α_3 , while other parts of the lid remains static. (B) Analysis of the C α displacements at the amino acid level supports the dynamics of lid α_3 between these two structures. In particular, Met75 has the highest C α rmsd (left) and moves approximately 1.6 Å toward the glucose moiety in the donor subsite. The main-chain displacements also affect the side-chain positional and conformational changes (right), such that the Met75 side-chain is shifted toward the active site with its sulfur closer to the donor subsite (inset in A).



β 4GT



GTA



hGYG1

```
hGYG1  ---MVSLEP---RMVYPPQP---KVLTP---C-R-KDVLVVTP---WLAPIVWEGETFNI DILNE-Q
b4GT   TACPEESPLLVG PMLIEFNIPV D LK LVEQQNPKV KLGGRYTPMD C I S P H K W A I I I P F R N R Q E H L K Y W ---LY-----YL

hGYG1    7           17           26           34           44           54           64           74
hGYG1  -SMTDQAF-VTLTTNDAYAKGALVL-GSSLLKQXRTRR--LVVLATPQVSDSMRKVLETVPDEVI MVDVLD SGDS AHL TL
GTA    FRLQNTTIGLTVFAIKKVVAF LKLF-LETAEKHFMVGH RVHYVFT-DQPAAVPRWT LGTGRQLSVLE VRAYKRWQD-VS
b4GT   HPILQRQ---QLDYGIYVW--INQ-AGESMFNR AKLLN---VGFKE-----

hGYG1    84           92           102          112          119          123
hGYG1  MKRPELGVTLLTKLHCW--SLTQVSKCVFMDADTLVLANIDDLFDREE-----LSAAP-----
GTA    MR R M E M -I S D F C E R R --F L S E V D Y L V C V D V D M E F R D H V G V E I L T P L ---FGTLHPGFY GSSREAF TYERRPQ
b4GT   ---ALKDYDYNCFVFS---DVDLIPMN-DHNTYRCFSQPRH- ISVA-MDKPG---P---SLP-

hGYG1    127          137          147          157          166          176          186          196
hGYG1  ---DPGWPD CFNSGVVFVYQPSVETYNQLLHLASEQGSFDDGGDQ-GILNTFFSSWATTDIRKXLPFIYNLSSISIYS
GTA    SQAYIIPKDEG-DFYYGGAGFFG-----GSVQEVORLTR-ACHQ
b4GT   ---Y-----VQYFGGWSA-----L S K Q Q F-----L S I N G F P N N Y W --G W G G E D D D -I Y N R L A F R G M S

hGYG1    205          213          222          232          242          252          259
hGYG1  YLPFAKVFVFG--ASAKVVVHF-LGRVKPWNYY-TYDPKTKSVKSEAHDPNMTHP E F I I L W W N I F ---TTNVLPLLO
GTA    AMMVDQANG--IEA-VWHD-ESH L N--KY-LLRHKPTKVL--SPEYLDW D Q Q -LLGWPAVLR--KLRFTA VPK
b4GT   VSRPNAVIGKTRMIRHSRDKKNEPNPQRFDRIAHATKETMLSDGL-NSLTVMVLEVQRYP LYTKI-TVDIGTFS
```

Fig. S7. Comparison of sugar donor-induced conformational changes in GT-A fold enzymes. The known structural alterations upon UDPG binding in GT-A type glycosyltransferases include bovine β 1,4-galactosyltransferase (b4GT, PDB ID codes 1O0R, 1FGX), human fucosylgalactoside α -N-acetylgalactosaminyltransferase (GTA, PDB ID codes 3IOH, 3IOJ) and human GYG1 (this study). The mobile region is coloured orange for the *apo* form, and blue for the *holo* (UDPG-bound) form. The core β -strands in the Rossmann-like GT-A fold are colored yellow, and UDPG ligands are in stick-and-ball representations. The comparison demonstrates that the conformational changes upon UDPG binding differ in the location, residue composition and extent/nature of movement among the enzymes, despite sharing a similar overall fold. In addition, the mobile regions are also unpredictable from their sequences, as shown in the sequence alignment (using hGYG1 residue numbering). The mobile regions are colored blue, and residues corresponding to the overlaid Rossmann-like fold are double-underlined.

