

Supplementary Information

Structure and mechanism of cancer-associated N-acetylglucosaminyltransferase-V

Nagae et al.

Supplementary Table 1 Nucleotide sequences of the primers used in this study

A restriction enzyme site (XhoI) is underlined.

Primer	Nucleotide sequence
pcDNA-IH/GnT-V T121 fwd	5'-CTACAGCTGTTCCCAGCTTG-3'
pcDNA-IH/GnT-V L741 rev	5'-AGGCTCGAGCTATAGGCAGTCTTTGCAGA-3'
N_del_fwd	5'- AGATCCGTCCTTTCCCTGGAAGTAC-3'
N_del_rev	5'- AATTCATTGGCGGAAATTCGTACAG-3'
Loop_change_fwd	5'-CTTCATGATTTCCCTGAGCTCAGCC-3'
Loop_change_rev	5'-ggcggcggcggcATTGTTGAGCTCATTTACATTGATA-3'
E280A (fwd and rev)	5'-CCTGGGACTCCTGACCAAGGCATCTGGATTAAAGATTGCAG-3' 5'-CTGCAATCTTAAATCCAGATGCCTTGGTCAGGAGTCCCAGG-3'
E287A (fwd and rev)	5'-ATCTGGATTAAAGATTGCAGCGACAGCTTTCAGTGGTGGCC-3' 5'-GGCCACCACTGAAAGCTGTCGCTGCAATCTTAAATCCAGAT-3'
E297A (fwd and rev)	5'-CAGTGGTGGCCCTCTTGGTGCATTAGTTCAATGGAGTGATT-3' 5'-AATCACTCCATTGAACTAATGCACCAAGAGGGCCACCACTG-3'
E429A (fwd and rev)	5'-CTTTCTGGGGTTTGTGGTTGCGCAGCACCTGAACTCCAGTG-3' 5'-CACTGGAGTTCAGGTGCTGCGCAACCACAAACCCAGAAAG-3'
E520A (fwd and rev)	5'-TGGACTTGGGTTCCCTTACGCGGGCCCAGCTCCCCTGGAAG-3' 5'-CTTCCAGGGGAGCTGGGCCCCGCGTAAGGGAACCCAAGTCCA-3'
E526A (fwd and rev)	5'-CGAGGGCCCAGCTCCCCTGGCAGCTATCGCAAATGGATGTG-3' 5'-CACATCCATTTGCGATAGCTGCCAGGGGAGCTGGGCCCTCG-3'

Supplementary Table 2

Interaction network between acceptor trisaccharide and mini-GnT-V

(i) Direct protein-glycan and (ii) water-mediated interactions of nitrogen and oxygen atoms whose distances are within 3.5 Å. The water-mediated interactions listed below are common in the two complexes in the asymmetric unit.

(i) Direct interaction

Trisaccharide	Distance (Complex A, B)	GnT-V
Man-2 O3	2.6, 2.9	K554 N ζ
Man-2 O3	3.4, -	K554 N ζ
GlcNAc-3 O3	2.9, 3.2	S379 N
GlcNAc-3 O3	3.5, 3.5	S379 O γ
GlcNAc-3 O4	2.5, 2.5	D378 O δ
GlcNAc-3 O5	2.8, 3.0	K554 N ζ
GlcNAc-3 O6	3.3, 3.0	K554 N ζ
GlcNAc-3 O7	3.1, -	S379 O γ

(ii) Water-mediated indirect interaction

Trisaccharide	Distance (A, B)	Water (A, B)	Distance (A, B)	GnT-V
GlcNAc-3 O3	2.8, 2.6	Wat69, 190	3.0, 2.8	L377 O
GlcNAc-3 O3	3.2, 3.3	Wat154, 52	3.0, 3.0	I353 N
GlcNAc-3 O4	2.8, 2.9		2.6, 2.7	L377 O
GlcNAc-3 O6	2.6, 2.6	Wat28, 172	2.5, 2.5	D378 O δ 1
			3.5, 3.4	D378 O δ 2
			3.2, 3.4	G381 N
			2.7, 2.9	G381 O

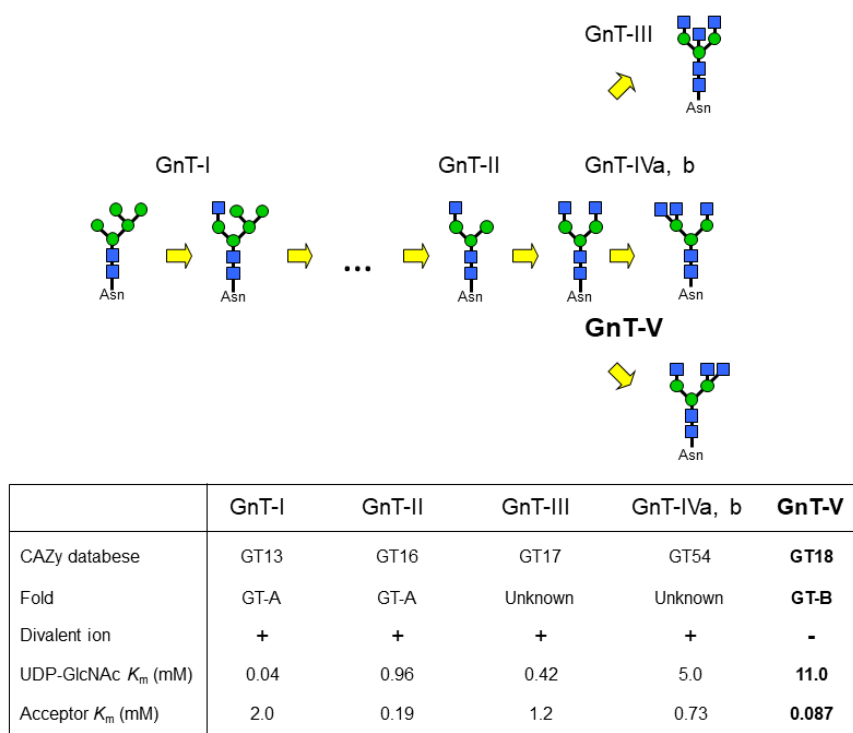
Supplementary Table 3

Gel filtration analysis of GnT-V luminal domain and mini-GnT-V E297A proteins.

The gel filtration column, Superdex75 HiLoad16/60 (GE Healthcare), was used for this experiment.

The molecular weights and elution points of standard proteins were as follows: Alcohol dehydrogenase (141kDa, 52.2mL), Bovine Serum Albumin (66.5kDa, 57.0mL) and Carbonic Anhydrase (29kDa, 72.4mL).

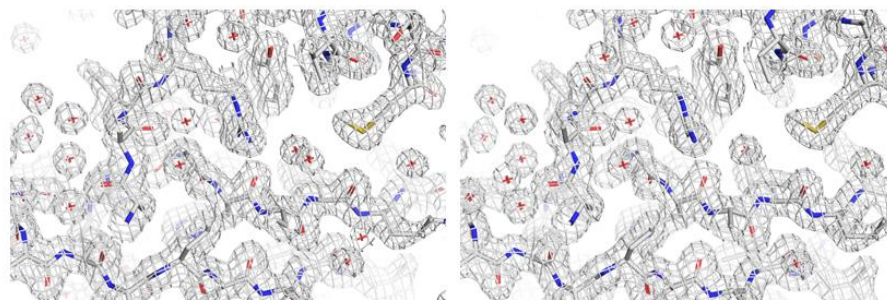
Proteins	Calculated MW (kDa)	Elution points (mL)	Estimated MW (kDa)
GnT-V luminal domain	71	57.9	81
Mini-GnT-V E297A	59	63.0	53.8



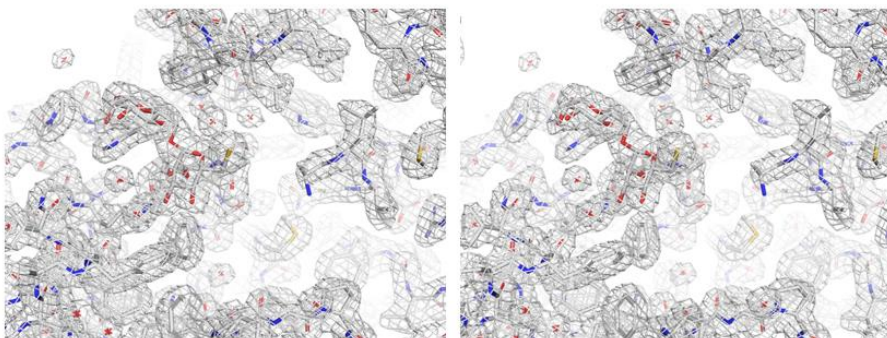
Supplementary Figure 1

Biosynthetic pathway of GlcNAc-branched N-glycans catalyzed by human GlcNAc transferases (GnT-I~V) is shown in upper panel. Enzymatic properties of human GlcNAc transferases are summarized in lower Table. This figure was modified from original reviews^{1,2} and recent report about crystallographic study of GnT-II³.

**GnT-V
Luminal
Domain
Apo-form**

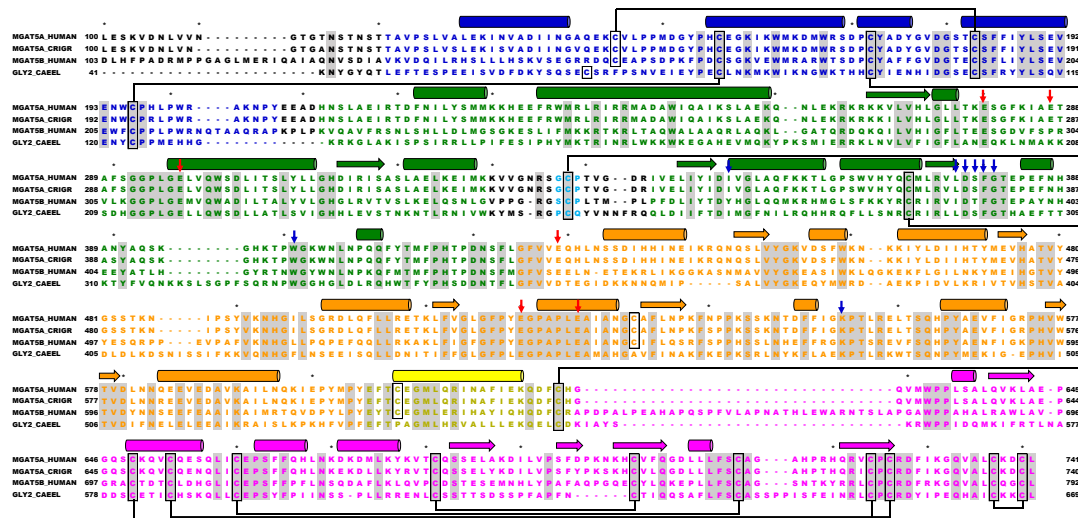


**Mini-GnT-V
E297A
complex**



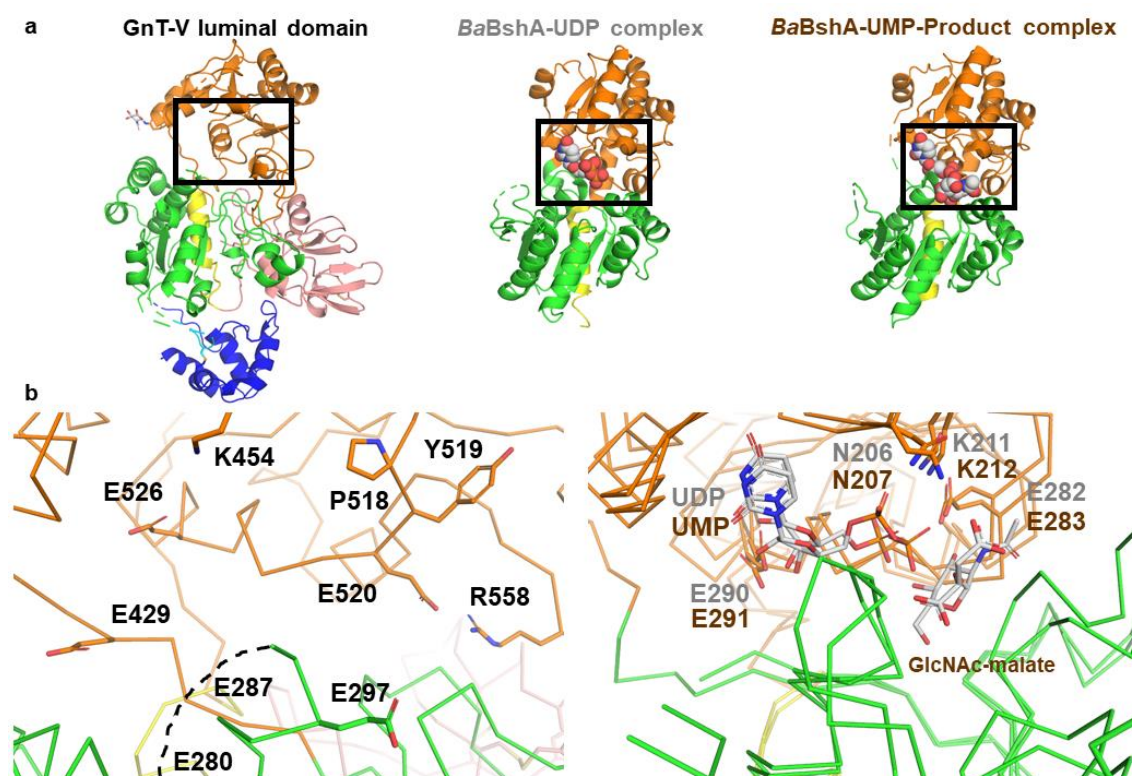
Supplementary Figure 2

The electron density map of refined structures. Stereo views of $2F_{\text{obs}} - F_{\text{calc}}$ electron density map contoured at 1.0 σ level of GnT-V luminal domain (upper panel) and Mini-GnT-V E297A inhibitor complex (lower panel) are shown.



Supplementary Figure 3

Amino acid sequence alignment of GnT-V and related GT18 family members. The alignment was performed with ClustalW server. Accession database numbers of each sequence are as follows: Q09328 (human GnT-V, MGAT5A_HUMAN), G3GZF3 (Chinese Hamster GnT-V, MGAT5A_CRIGR), Q3V5L5 (human GnT-IX, MGAT5B_HUMAN), and Q9NDH7 (*C.elegans* Gly-2, GLY2_CAEEL). Cysteine residues are enclosed in boxes and disulfide bond pattern is indicated. Secondary structural elements of human GnT-V are indicated above the sequence. Six glutamates subjected to mutational experiments, are indicated with red arrows. Amino acid residues which interact with acceptor trisaccharide are indicated with blue arrows. The amino acid sequence identities with human GnT-V are 97% (Chinese Hamster GnT-V), 40% (human GnT-IX) and 28% (*C.elegans* Gly-2).

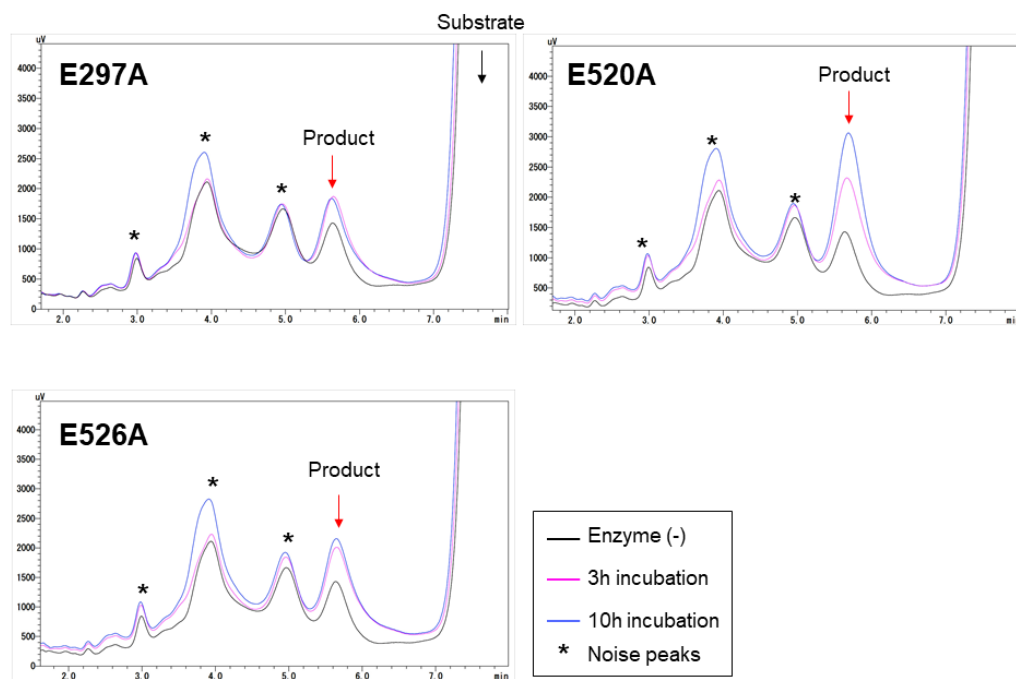


Supplementary Figure 4

Structural comparison between GnT-V luminal domain and bacterial glycosyltransferase, *BaBshA*

(a) Structural comparison between GnT-V luminal domain (left panel), bacterial glycosyltransferase, *BaBshA*-UDP complex (middle panel, PDB code: 3MBO⁴) and *BaBshA*-UMP-GlcNAc-malate complex (right panel, PDB code: 5D00⁵). Protein and ligand molecules are shown in ribbon and sphere models, respectively. Three structures are depicted from the same view angles. Ligand binding sites are highlighted in black boxes.

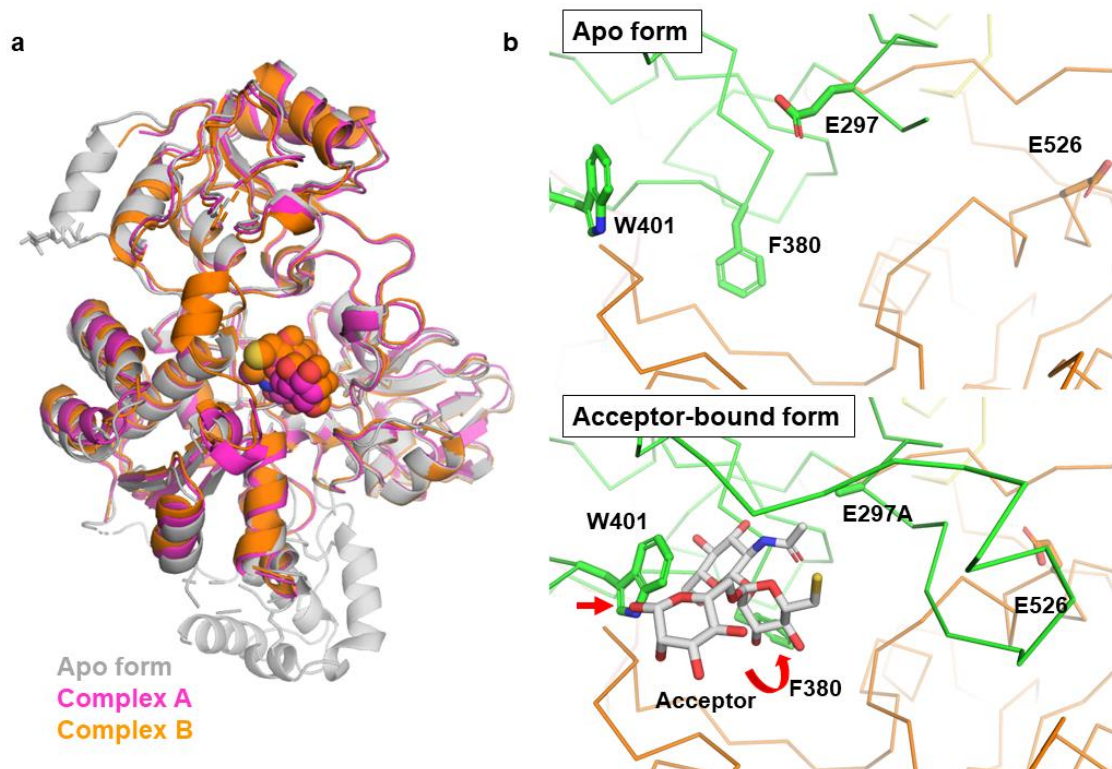
(b) Structural comparison of donor and product binding sites of *BaBshA* (right panel) and corresponding region of GnT-V luminal domain (left panel). Two structures of *BaBshA* complexes are superimposed in right panel. Polar residues which interact with UDP in *BaBshA* and corresponding residues in GnT-V are shown in rod models. Also, six glutamates applied to mutational experiments are shown in rod models.



Supplementary Figure 5

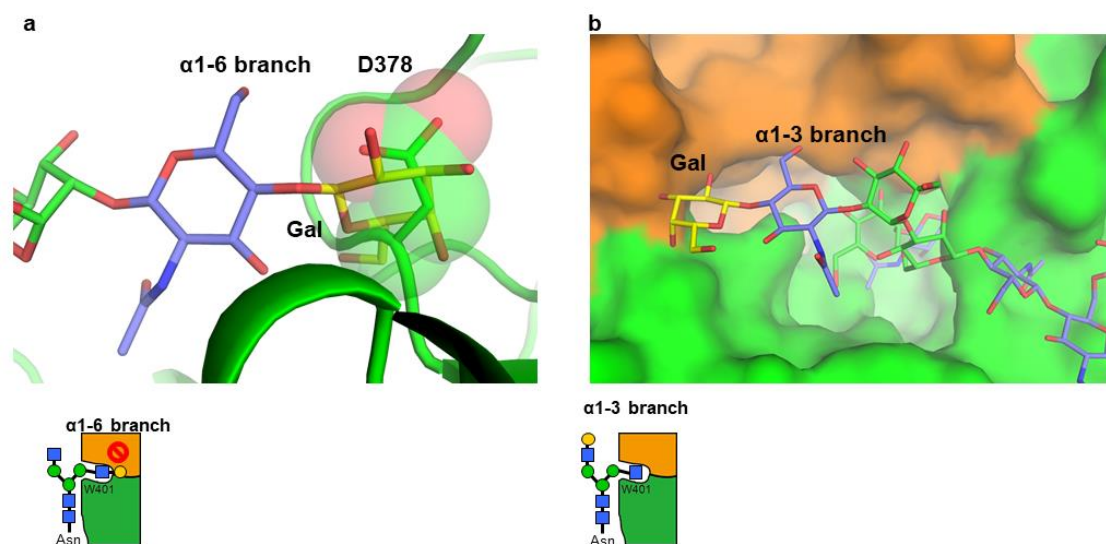
Enzymatic activity of three GnT-V mutants (E297A, E520A and E526A) with prolonged reaction time.

Elution profiles of the substrate and product after the incubation of each mutant with the substrate for three (pink) or ten (blue) hours. The positions of the product were indicated.



Supplementary Figure 6

- (a)** Structural superposition of apo form (grey) and two complexes (A: magenta, B: orange) in the asymmetric unit. Proteins and ligands are shown in ribbon and sphere models, respectively.
- (b)** Structural comparison between apo form (upper panel) and acceptor-bound form (lower panel) focusing on the catalytic center. Sugar residues and amino acid residues involved in catalytic reaction are shown in rod models and labeled. Observed structural differences upon acceptor binding are indicated with red arrows.



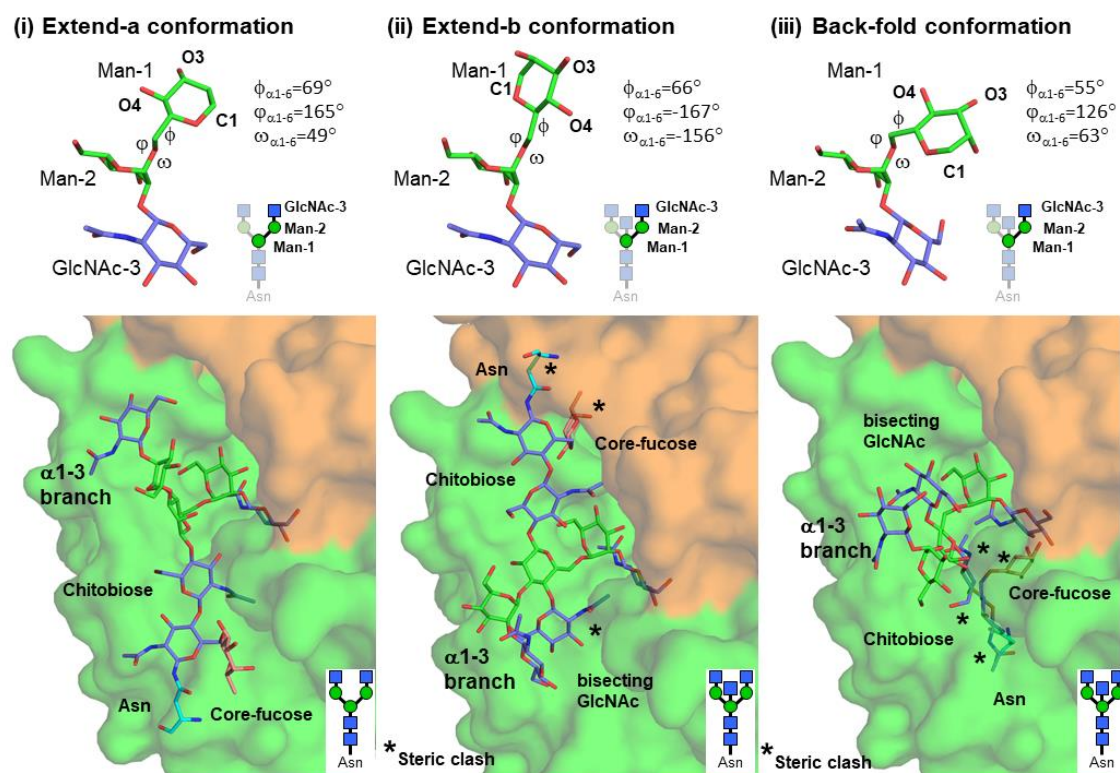
Supplementary Figure 7

(a) Docking models of α 1-3 branch of biantennary glycan onto α 1-6 branch. As described in main text, β -mannose causes steric clash with W401. In addition, chitobiose also causes heavy clash with S281-G282.

(b) Docking models of galactosylated N-glycans with GnT-V

Left panel: Galactose extension at α 1-6 branch. Structural superposition of galactosylated α 1-6 branch onto the crystal structure of mini-GnT-V trisaccharide complex. The galactose sterically clashes with D378 of mini-GnT-V.

Right panel: Galactose extension at α 1-3 branch. N-glycan is shown in rod model. Mini-GnT-V is shown in surface model.



Supplementary Figure 8

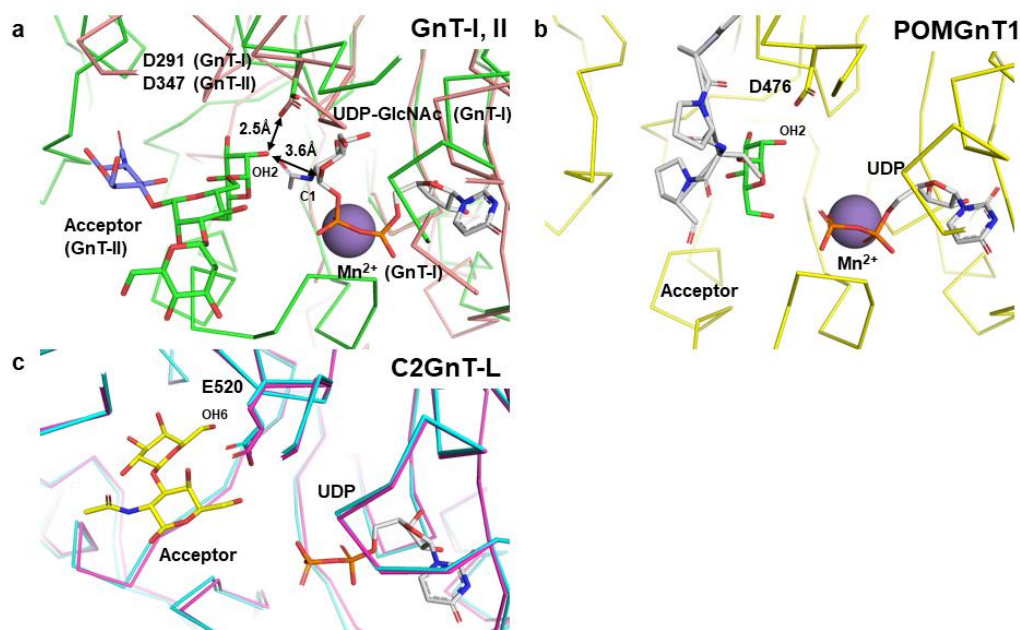
Conformation selective inhibition mechanism of GnT-V. The dihedral angles ($\phi_{\alpha1-6}$, $\phi_{\alpha1-6}$, $\omega_{\alpha1-6}$) of $\alpha1-6$ linkages between Man-1 (β -mannose) and Man-2 affect global conformation of N-glycans. Three different conformations of trisaccharide unit (GlcNAc β 1-2Man α 1-6Man) are shown in upper panels and docking models of three whole N-glycans onto mini-GnT-V are shown in lower panels.

(i) Trisaccharide unit of extend-a conformation ($\phi_{\alpha1-6}=69^\circ$, $\phi_{\alpha1-6}=165^\circ$, $\omega_{\alpha1-6}=49^\circ$) of biantennary N-glycan observed in crystal structure is shown in upper panel. Docking model of full N-glycan fits well into the cavity of mini-GnT-V is shown in lower panel.

(ii) Trisaccharide unit of extend-b conformation ($\phi_{\alpha1-6}=66^\circ$, $\phi_{\alpha1-6}=-167^\circ$, $\omega_{\alpha1-6}=-156^\circ$) of bisected N-glycan retrieved from the crystal structure of bisected N-glycan in complex with murine DCIR2 (PDB code: 3VYK, ⁶) is shown in upper panel. Docking model of bisected glycan complex is shown in lower panel. Compared with extend-a conformation, relative orientation of stem asparagine is flipped. It causes steric clash with mini-GnT-V. Steric clashes with mini-GnT-V are indicated with asterisks.

(iii) Trisaccharide unit of back-fold conformation ($\phi_{\alpha1-6}=55^\circ$, $\phi_{\alpha1-6}=126^\circ$, $\omega_{\alpha1-6}=63^\circ$) of bisected N-glycan extracted from the crystal structure of bisected N-glycan in complex with PHA-E (PDB code: 5AV7, ⁷) is shown in upper panel. Docking model with mini-GnT-V is displayed in lower

panel. Compared with extend-a conformation, chitobiose and core asparagine residues are deeply buried inside mini-GnT-V. Steric clashes are also indicated with black asterisks.



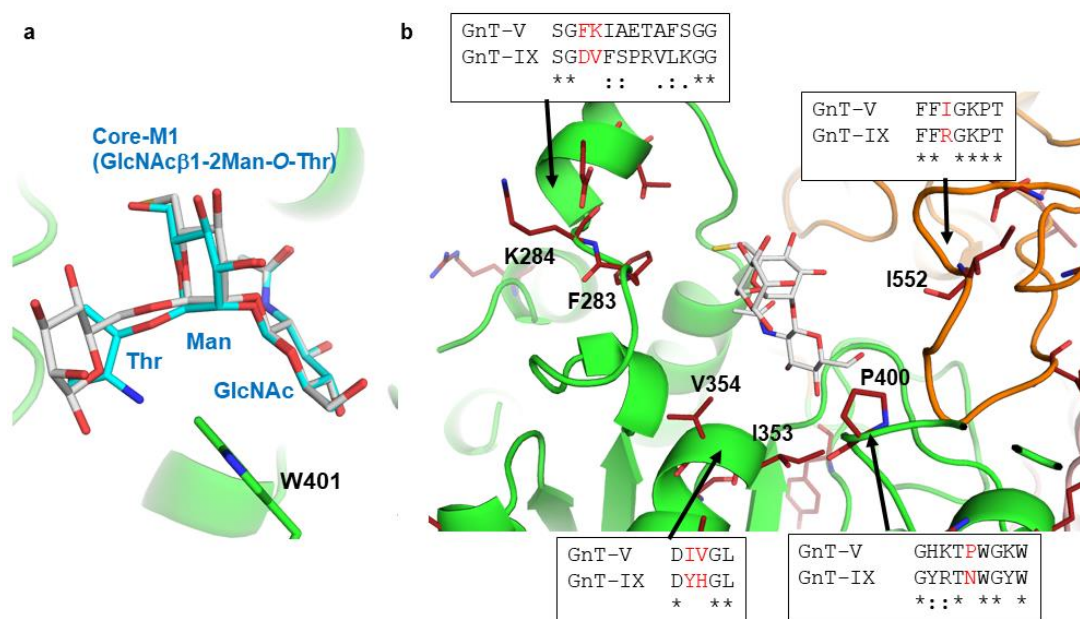
Supplementary Figure 9

Close-up views of substrate binding sites of other mammalian GlcNAc transferases which have GT-A folds.

(a) Close-up view of GnT-II-acceptor complex (PDB code: 5VCS, ³) posed onto GnT-I-Mn²⁺-UDP-GlcNAc complex (PDB code: 1FOA, ⁸). The catalytic base residues are located at same positions. The distance between OH2 of acceptor in GnT-II and D347 is 2.5 Å, whereas the distance between OH2 of acceptor in GnT-II and C1 of UDP-GlcNAc in GnT-I is 3.6 Å.

(b) Close-up view of POMGnT1 in complex with UDP, Mn²⁺ ion and acceptor Glycopeptide (PDB code: 5GGI, ⁹)

(c) Close-up view of C2GnT-L in complex with acceptor disaccharide (PDB code: 2GAM, ¹⁰) posed onto UDP complex (PDB code: 3OTK, ¹¹).

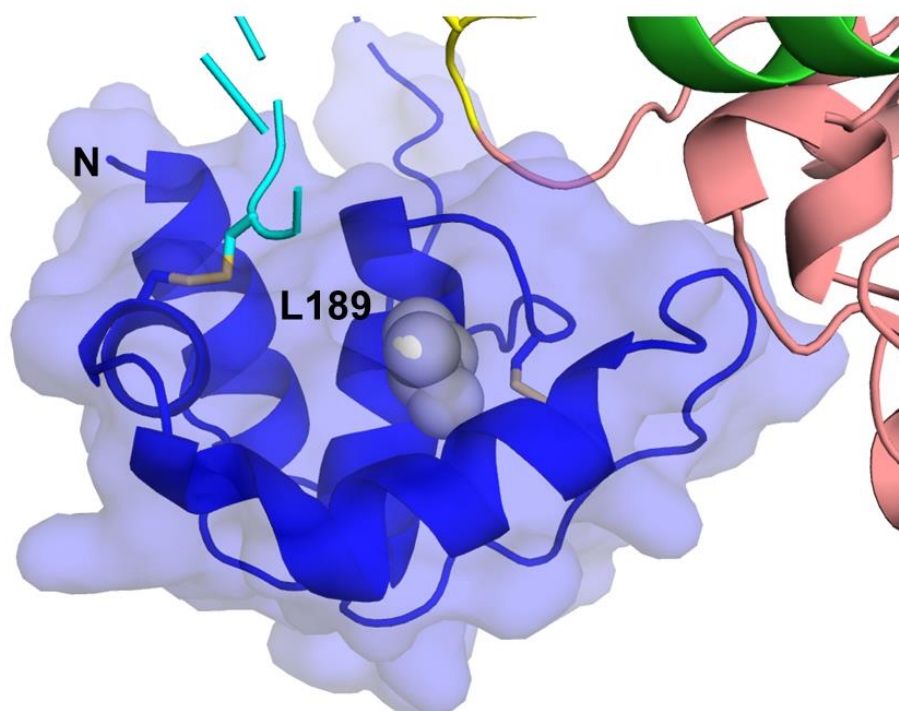


Supplementary Figure 10

Structure-based sequence comparison with GnT-IX (GnT-Vb)

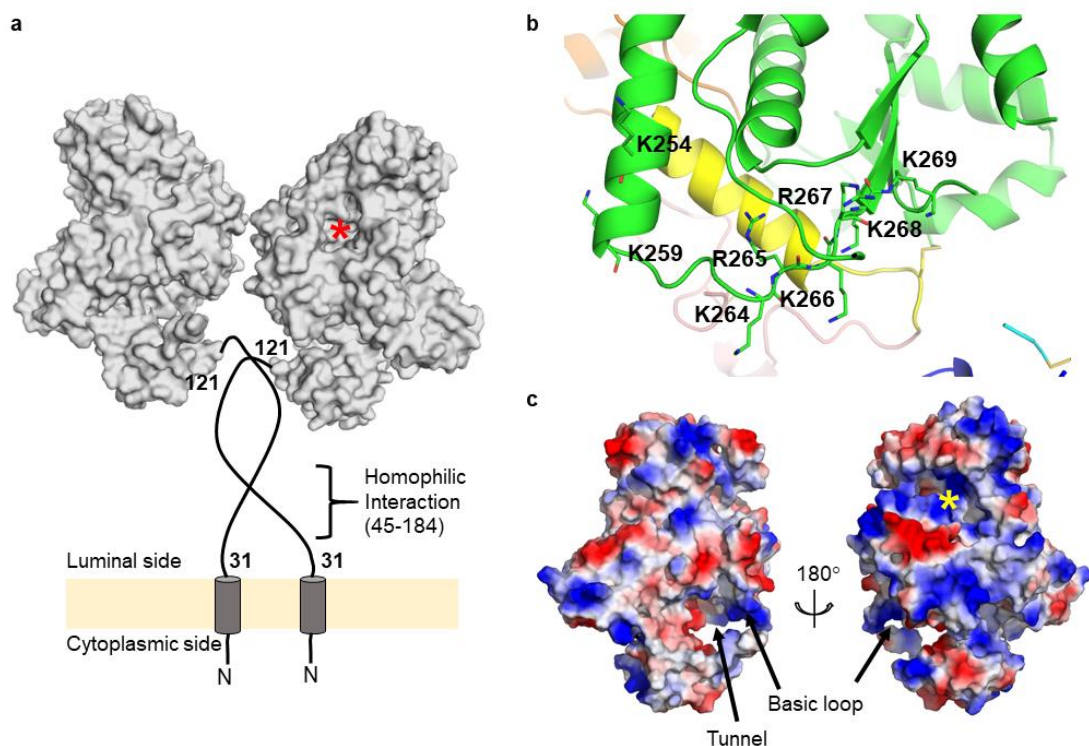
(a) Hypothetical structure of Core-M1 glycan (GlcNAc β 1-2Man-O-Thr, cyan) was superposed onto mini-GnT-V trisaccharide (GlcNAc β 1-2Man α 1-6Man) complex. Sugar residues and W401 are shown in rod model.

(b) Structure-based sequence comparison between GnT-V and IX. Incompatible amino acid residues are shown in rod model. Seven amino acid residues which are different in GnT-IX are shown in rod models and labeled. Partial sequence alignments around acceptor binding site are indicated with boxes.



Supplementary Figure 11

Close-up view of N-terminal domain. N-terminal domain is shown in ribbon and semi-transparent surface models. The replaced residue in Lec4A mutant (L189R) is highlighted in sphere model.



Supplementary Figure 12

(a) Hypothetical dimer formation via homophilic interaction region. The catalytic center is indicated with red asterisk.

(b) Close-up view of positively charged loop (K264-K269) and adjacent lysine residues (K254 and K259) in middle domain 1. Positively charged residues are shown in rod models.

(c) Electrostatic surface representation of GnT-V luminal domain. The position of catalytic center is indicated with asterisk. The position of basic loop and adjacent tunnel are also labeled.

Supplementary References

1. Varki, A. *Essentials of glycobiology*, xxix, 823 pages (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2017).
2. Dennis, J.W., Nabi, I.R. & Demetriou, M. Metabolism, cell surface organization, and disease. *Cell* **139**, 1229-41 (2009).
3. Kadirvelraj, R. et al. Human N-acetylglucosaminyltransferase II substrate recognition uses a modular architecture that includes a convergent exosite. *Proc Natl Acad Sci U S A* **115**, 4637-4642 (2018).
4. Parsonage, D. et al. Characterization of the N-acetyl-alpha-D-glucosaminyl l-malate synthase and deacetylase functions for bacillithiol biosynthesis in *Bacillus anthracis*. *Biochemistry* **49**, 8398-414 (2010).
5. Winchell, K.R. et al. A Structural, Functional, and Computational Analysis of BshA, the First Enzyme in the Bacillithiol Biosynthesis Pathway. *Biochemistry* **55**, 4654-65 (2016).
6. Nagae, M. et al. Recognition of bisecting N-acetylglucosamine: structural basis for asymmetric interaction with the mouse lectin dendritic cell inhibitory receptor 2. *J Biol Chem* **288**, 33598-610 (2013).
7. Nagae, M. et al. Atomic visualization of a flipped-back conformation of bisected glycans bound to specific lectins. *Sci Rep* **6**, 22973 (2016).
8. Unligil, U.M. et al. X-ray crystal structure of rabbit N-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily. *EMBO J* **19**, 5269-80 (2000).
9. Kuwabara, N. et al. Carbohydrate-binding domain of the POMGnT1 stem region modulates O-mannosylation sites of alpha-dystroglycan. *Proc Natl Acad Sci U S A* **113**, 9280-5 (2016).
10. Pak, J.E. et al. X-ray crystal structure of leukocyte type core 2 beta1,6-N-acetylglucosaminyltransferase. Evidence for a convergence of metal ion-independent glycosyltransferase mechanism. *J Biol Chem* **281**, 26693-701 (2006).
11. Pak, J.E., Satkunarajah, M., Seetharaman, J. & Rini, J.M. Structural and mechanistic characterization of leukocyte-type core 2 beta1,6-N-acetylglucosaminyltransferase: a metal-ion-independent GT-A glycosyltransferase. *J Mol Biol* **414**, 798-811 (2011).