Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Nat Chem Biol. Author manuscript; available in PMC 2023 December 18.

Published in final edited form as:

Nat Chem Biol. 2023 August ; 19(8): 1022–1030. doi:10.1038/s41589-023-01345-y.

Structural basis for Lewis antigen synthesis by the α1,3-fucosyltransferase FUT9

Renuka Kadirvelraj[1,5], **Bhargavi M. Boruah**[2,5], **Shuo Wang**[2,5], **Digantkumar Chapla**[2], **Chin Huang**[1,2], **Annapoorani Ramiah**[2], **Kieran L. Hudson**[3], **Anthony R. Prudden**[4], **Geert-Jan Boons**[2,4], **Stephen G. Withers**[3], **Zachary A. Wood**[1,✉], **Kelley W. Moremen**[1,2,✉]

[1]Department of Biochemistry & Molecular Biology, University of Georgia, Athens, GA, USA.

[2]Complex Carbohydrate Research Center, University of Georgia, Athens, GA, USA.

[3]Department of Biochemistry and Molecular Biology, Department of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada.

[4]Department of Chemistry, University of Georgia, Athens, GA, USA.

[5]These authors contributed equally: Renuka Kadirvelraj, Bhargavi M. Boruah, Shuo Wang.

Abstract

Mammalian cell surface and secreted glycoproteins exhibit remarkable glycan structural diversity that contributes to numerous physiological and pathogenic interactions. Terminal glycan structures include Lewis antigens synthesized by a collection of α1,3/4-fucosyltransferases (CAZy GT10 family). At present, the only available crystallographic structure of a GT10 member is that of the *Helicobacter pylori* α1,3-fucosyltransferase, but mammalian GT10 fucosyltransferases are distinct in sequence and substrate specificity compared with the bacterial enzyme. Here, we determined crystal structures of human FUT9, an α1,3-fucosyltransferase that generates Lewis^x and Lewis^y antigens, in complex with GDP, acceptor glycans, and as a FUT9–donor analog–acceptor Michaelis complex. The structures reveal substrate specificity determinants and allow

✉ **Correspondence and requests for materials** should be addressed to Zachary A. Wood or Kelley W. Moremen. zaw@uga.edu; moremen@uga.edu.

Author contributions

K.W.M. and Z.A.W. formulated the project. B.M.B., S.W. and D.C. expressed and purified recombinant proteins. B.M.B., S.W., D.C. and R.K. performed structural studies. B.M.B., D.C. and C.H. performed enzyme kinetic analysis. A.R. generated expression constructs and site-directed mutants. K.L.H. and S.G.W. synthesized GDP-CF3-Fuc. A.R.P. and G.-J.B. synthesized glycan substrates. R.K., B.M.B., Z.A.W. and K.W.M. wrote the paper.

Competing interests

K.W.M. is the President of Glyco Expression Technologies, a biotechnology spinout that focuses on production and commercialization of recombinant glycosyltransferases, which could conceivably profit from the results described herein. The remaining authors have no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41589-023-01345-y.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41589-023-01345-y.

Peer review information *Nature Chemical Biology* thanks Rhys Grinter, Robert Sackstein and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

prediction of a catalytic model supported by kinetic analyses of numerous active site mutants. Comparisons with other GT10 fucosyltransferases and GT-B fold glycosyltransferases provide evidence for modular evolution of donor- and acceptor-binding sites and specificity f or Lewis antigen synthesis among mammalian GT1 0 fucosyltransferases.

Cell surface and secreted glycoproteins and glycolipids contain glycan structures that play critical roles as an interface with the extracellular environment[1,2]. Terminal glycan structures act as ligands in numerous biological recognition events that influence cell signaling, cell adhesion, modulation of circulating glycoprotein and receptor half-lives, and host–pathogen interactions, among many other roles[1–3].

Terminal Lewis antigen epitopes on glycoproteins and glycolipids consist of a subset of histo-blood group antigens (Lewis^a and Lewis^b antigens: Le^a and Le^b) and cell surface structures on epithelial cells and numerous other cell types (Le^x and Le^y)[4,5]. Sialylated and sulfated Lewis antigens also play critical roles in development, neutrophil transepithelial migration, immunoregulation and cancer[6–8]. Lewis antigen synthesis occurs through addition of fucose (Fuc) residues in α1,4-Fuc linkage to the subterminal GlcNAc residue in type 1 LacNAc (Gal-β1,3-GlcNAc-) structures to form Le^a epitopes, or in α1,3-Fuc linkage to GlcNAc residues in type 2 LacNAc (Gal-β1,4-GlcNAc-) chains to form Le^x epitopes (Fig. 1)[4,9]. Additional complexity in Lewis antigen structures comes through elaboration of LacNAc units by previous α1,2-fucosylation of the Gal residue (H antigen synthesis) or sialylation (sialyl-LacNAc) and/or sulfation to form the corresponding modified Lewis antigen derivatives[4,5].

Synthesis of the complex series of Lewis antigen epitopes is achieved by a family of fucosyltransferases belonging to the CAZy GT10 glycosyltransferase family[4,10–13]. These enzymes are distinct in structure and specificity compared with other mammalian fucosyltransferases found in CAZy GT11 (FUT1 and FUT2, involved in H antigen synthesis[14–16]), GT23 (FUT8, catalyzing N-glycan core fucosylation[17,18]), GT65 and GT68 (POFUT1 and POFUT2, protein domain-specific Ser/Thr O-fucosyltransferases[4,19]) and additional nonmammalian fucosyltransferases (GT37 (ref. 20) and GT74 (ref. 21)). GT10 fucosyltransferases are widely found in bacteria, eukaryotes and viruses[10] and all enzymes tested so far employ GDP-Fuc as donor for transfer to various glycan acceptors. Humans have eight GT10 isoforms (FUT3–7, FUT9–11) each with distinctive acceptor specificities that broadly fall into three clades reflecting their evolution from a common ancestor[13]. FUT4, FUT7 and FUT9 are α1,3-fucosyltransferases that act on type 2 LacNAc acceptors to generate Le$^{x/y}$ and derivative structures, but can be distinguished in their abilities to form sialyl-Le^x epitopes[22–24]. FUT3, FUT5 and FUT6 arose from a recent gene expansion in primates but are more heterogeneous in specificity[25–27]. FUT3 and FUT5 are the only enzymes that can modify both type 1 and type 2 LacNAc chains as bifunctional α1,3/4-fucosyltransferases and are responsible for Le$^{a/b}$ histo-blood group antigen synthesis[14,25,27]. In contrast, FUT6 synthesizes only α1,3-Fuc linkages on type 2 chains to generate Le$^{x/y}$ and sialyl-Le^x structures[26]. By comparison, FUT10 and FUT11 diverged early in the GT10 lineage and apparently can synthesize α1,3-Fuc linkages on N-glycan core residues in

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

vitro[28]. Thus, each mammalian fucosyltransferase has a unique acceptor specificity toward complex glycan structures that distinguishes it in Lewis antigen synthesis.

So far, only one GT10 structure has been solved, the α1,3-fucosyltransferase from the human pathogen *Helicobacter pylori* (HpFucT), as an apoprotein and in complex with GDP or GDP-Fuc[29,30]. This enzyme forms Le^x and Le^y epitope structures on bacterial lipopolysaccharides that mimic mammalian host cell surface Lewis antigens[29]. The enzyme exhibits a GT-B glycosyltransferase fold[31] with one Rossmann-like domain involved in sugar donor interactions, and a second Rossman-like domain presumed to be involved in acceptor interactions. By comparison, human GT10 isoforms share <21% sequence identity with HpFucT (Extended Data Fig. 1b) and have substantial differences in substrate specificity[22–29].

Here we present the structural basis for Lewis antigen synthesis by human FUT9. The catalytic domain structure was determined both ligand-free and in complex with GDP, acceptor ligands lacto-*N*-neotetraose (LNnT) and H-type 2 (with and without GDP), and in complex with a poorly hydrolyzable GDP-6,6,6-trifluoro-β-L-fucose (GDP-CF_3-Fuc) donor in combination with an H-type 2 acceptor as a Michaelis complex. To further map the details of donor and acceptor interactions and catalysis, we performed kinetic analysis on a series of active site mutants to validate the structural model and compared the FUT9 structure with the HpFucT and bifunctional FUT3 structure. These data provide a framework for understanding the structural basis for Lewis antigen biosynthesis and the varied substrate specificities for the other mammalian GT10 fucosyltransferases that create a diverse collection of modified Lewis antigen structures.

Results

Human FUT9 structure

The soluble catalytic domain (residues 39–359) of human FUT9 was expressed in HEK293 cells as a secreted, N-terminally His-tagged GFP-Fc fusion protein[32,33] (Extended Data Fig. 2a). Following purification, the fusion tags and glycans were cleaved with TEV protease and EndoF1, respectively (Extended Data Fig. 2b). Size-exclusion chromatography–multiangle light scattering (SEC–MALS) identified a single broad peak eluting with a predicted molecular mass of ~77 kDa, indicating that the 37.8-kDa enzyme forms a dimer in solution (Extended Data Fig. 2c).

The crystal structure of FUT9 was solved using single-wavelength anomalous dispersion of a CsI heavy atom derivative (FUT9–HA; Supplementary Table 1), and the resolution was then extended to 1.1 Å using a native dataset (FUT9). The following FUT9 complexes were also crystallized and solved in the same $I2_13$ crystal lattice: FUT9–GDP, FUT9–LNnT, FUT9–H-type 2, FUT9–GDP–LNnT, FUT9–GDP–H-type 2 and FUT9–GDP-CF_3-Fuc–H-type 2 (a Michaelis complex with a fluorinated sugar donor, described below) (Supplementary Table 1). The high-resolution ligand-free crystal structure revealed a single chain in the asymmetric unit (Fig. 2a), but the application of crystallographic symmetry produces a dimer (Fig. 2b,c) with an interface that buries ~1,077 $Å^2$ of surface area and is consistent with the SEC–MALS results (Extended Data Fig. 2c). PISA[34] analysis predicts

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

that the dimer is stable, with a favorable solvation free energy gain ($\Delta^i G$) of −11.7 kcal mol^{-1} and a P value of 0.175, suggesting a specific interaction surface. The N-terminal 'stem region' of the crystal structure is missing 24 amino acids that are disordered. In the FUT9 dimer, both N termini are located on the same face of the complex, suggesting an arrangement where the dimer is anchored to the Golgi membrane via the N-terminal transmembrane segments (removed in the design of the expression construct) with the catalytic domain facing into the Golgi lumen (Fig. 2c).

Three N-glycosylation sites are predicted in the FUT9 sequence at Asn62, Asn101 and Asn153. The Asn62 residue could not be modeled since it resides in the disordered N-terminal stem region. Asn101 faces the solvent and the predicted GlcNAc residue attached to the side chain could not be resolved in the electron density. In contrast, a tetrasaccharide (Manα1,3-Manβ1,4-GlcNAcβ1,4-GlcNAcβ-) is clearly resolved attached to Asn153 (Fig. 2a–c). The glycan extends from a deep cleft formed in the dimer interface, which probably explains why it was protected from EndoF1 digestion. This glycan site is conserved in all human GT10 enzymes except FUT7 (Extended Data Fig. 1).

FUT9 adopts the characteristic structure of GT-B fold enzymes[31,35], consisting of an N-terminal acceptor- and C-terminal donor-binding domain. (Fig. 2a). The donor domain (residues 169–326) is present as an insertion between the β5 strand and α12 helix of the acceptor domain (residues 63–168, 327–359). A Dali search identified HpFucT (PDB 2NZX (ref. 29)) as the closest structural homolog, which is not surprising since both enzymes belong to CAZy GT10 and employ GDP-Fuc as the donor to synthesize α1,3-Fuc linkages on type 2 LacNAc units, albeit with distinct acceptor specificities (Supplementary Table 2)[24,29]. The donor domains of both enzymes contain a core Rossmann-like fold and superimpose 127 Cα atoms with a root mean squared deviation (r.m.s.d.) of 1.7 Å and a sequence identity of 27% (Extended Data Figs. 1 and 3). The acceptor-binding domains are more divergent and superimpose 80 Cα atoms with a r.m.s.d. of 3.7 Å and a sequence identity of 13.7%. The acceptor domains of both enzymes conserve a core fold consisting of a six-stranded parallel β-sheet, but the position and identity of the other secondary structural elements inserted into the fold are more variable. The Dali[36] search showed that the FUT9 acceptor domain belongs to the Toll/interleukin-1 receptor (TIR) fold family[37], which resembles a Rossmann-like fold. The insertion of the donor domain between the β5 strand and α12 helix (FUT9 numbering) of the acceptor fold is conserved in both enzymes, but the orientation of the α12 helix has shifted substantially between the proteins, which is reflected in the structural divergence of the associated linker elements between the acceptor and donor domains (Extended Data Fig. 3). Similar to FUT9, full-length HpFucT is also reported to form a dimer using an N-terminal coiled-coil motif combined with a proposed dimerization interface based on crystal contacts[29]. However, the reported dimerization interface is distinct from that observed in FUT9 (Extended Data Fig. 4). Other differences include three disulfides in FUT9 (Cys82–Cys335, Cys91–Cys338 and Cys190–Cys238), which are not conserved in HpFucT (Extended Data Fig. 1). In addition, FUT9 contains a *cis*-Phe329 (ω angle of ~13°) that is also not conserved in HpFucT (Extended Data Fig. 5).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The Dali search also identified structural similarities between FUT9 and other inverting GT-B fold enzymes from CAZy families 1, 9, 18, 28, 63, 70 and 90 (Supplementary Table 2). However, Dali searches revealed no noteworthy structural similarities to the GT-B fold fucosyltransferases from CAZy families GT23, GT37, GT65 or GT68, beyond the general features of the C-terminal Rossmann fold donor domain. This structural divergence is not surprising since sequence analysis partitions the GT10 α1,3/4-fucosyltransferases into a distinct clade relative to the α1,2-, α1,6- and protein-*O*-fucosyltransferases[13,38].

FUT9 acceptor interactions

The structure of FUT9 was also solved in complex with either the LNnT acceptor (Galβ1,4-GlcNAcβ1,3-Galβ1,4-Glc) (FUT9–LNnT) or the H-type 2 acceptor (Fucα1,2-Galβ1,4-GlcNAc) (FUT9–H-type 2), as well as the corresponding ternary complexes containing GDP (FUT9–GDP–LNnT and FUT9–GDP–H-type 2, respectively) (Fig. 3 and Supplementary Table 1). Both acceptors are bound in a surface pocket located in the cleft between the acceptor- and donor-binding domains. The acceptor-binding pocket is formed by four loops originating from the N-terminal acceptor domain: $Loop_{69–81}$, $Loop_{136–153}$, $Loop_{157–169}$ and $Loop_{326–330}$ with the latter two loops forming the linkers that connect the two domains (Fig. 3a,d). For the LNnT complexes, the LacNAc unit of the tetrasaccharide acceptor at the nonreducing end is well resolved, but the following Gal residue is solvent exposed and weakly ordered, and the terminal Glc residue is completely disordered and was not modeled (Fig. 3b). For the H-type 2 complexes, the electron density was clearly resolved for the entire trisaccharide (Fig. 3e). The electron density for the GlcNAc residue at the reducing end supports a mixture of both the α- and β-anomers (Fig. 3e, inset). The α1,2-Fuc residue faces the solvent and is ordered by packing against the carboxamide of Gln75 (Fig. 3e). The additional packing interaction probably explains the ordering of the acceptor and may explain why FUT9 displays a slight preference for the H-type 2 acceptor (4.2-fold higher catalytic efficiency (k_{cat}/K_m) relative to the LNnT acceptor) (Supplementary Table 3). The conformation and binding interactions with the LacNAc unit of LNnT are essentially identical to those observed in the H-type 2 complexes, except where indicated below. Loop residues Phe73, Leu136, His141, Tyr168, Phe329 and Trp330 form a complementary surface for the LacNAc unit of the acceptor, while the Gln75 Nε donates a hydrogen to the O2 hydroxyl of the Gal (Fig. 3b,c). Phe73 supports and positions the GlcNAc residue in the active site, and the carboxylate of Glu137 forms hydrogen bonds with both the GlcNAc O3 and Gal O6 hydroxyl groups (Fig. 3c,f). This interaction identifies Glu137 as the putative catalytic base for deprotonating the GlcNAc O3 hydroxyl to act as the nucleophile in the transferase reaction. Consistent with this interpretation, the E137Q substitution showed no detectable enzyme activity, while the E137A substitution reduced k_{cat}/K_m by 2,400- and 2,870-fold against LNnT and H-type 2 acceptors, respectively (Supplementary Table 3). Similarly, the F73A substitution ($Loop_{69–81}$) shows no detectable activity, which indicates its importance in positioning and stabilizing the GlcNAc O3 nucleophile. The Q75A, E137A, H141A and F329A substitutions reduced k_{cat}/K_m for the H-type 2 acceptor by 68-, 2,870-, 97- and 421-fold, respectively, demonstrating the importance of the $Loop_{69–81}$, $Loop_{136–153}$, $Loop_{157–169}$ and $Loop_{326–330}$ binding pocket structural elements in acceptor specificity (Supplementary Table 3 and Fig. 3b,e).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

There is only a slight difference between the interactions with the LacNAc unit in the LNnT and H-type 2 acceptors (Fig. 3b,c,e,f). In LNnT, Gln75 donates a hydrogen to the Gal O2, but in the H-type 2 acceptor the Gal O2 hydroxyl has been extended by an α1,2-linked Fuc residue. While Gln75 Nε2 still contributes a hydrogen to the O2 of the fucose, the distance has increased from 3.0 Å to 3.4 Å, indicating a very weak hydrogen bond (it is modeled as a van der Waals/electrostatic interaction in the Ligplot (Fig. 3f)). A comparison of H-type 2 and LNnT complexes with ligand-free FUT9 reveals no notable conformational changes on ligand binding (r.m.s.d. values of ~0.13 Å). This suggests that either FUT9 does not undergo a conformational change on substrate binding, or, because the crystals are isomorphous, the lattice may have selected for the ligand-free enzyme in the acceptor-bound conformation.

In contrast to FUT9, the HpFucT acceptor substrate is a lipopolysaccharide. Thus, it is not surprising that the structure of the loop elements that form the HpFucT acceptor-binding site and the residues involved in acceptor specificity are quite different. While there is no available structure of the acceptor complex of HpFucT, the similar positioning of the sugar donor and catalytic bases (Glu137 in FUT9 versus Glu95 in HpFucT[29]) between these two enzymes suggests that the nucleophile of the *H. pylori* lipopolysaccharide acceptor (GlcNAc O3 hydroxyl) will be similarly positioned in the corresponding active sites (Extended Data Fig. 3i).

Donor sugar binding

Efforts to obtain an intact FUT9–GDP-Fuc complex through ligand soaking or cocrystallization yielded electron density for only the GDP moiety, which showed that the crystallized enzyme was active and hydrolyzing the donor sugar on binding (Extended Data Fig. 6). To obtain a Michaelis complex, we used a GDP-Fuc derivative containing a trifluoromethyl group at the C6 position (GDP-6,6,6-trifluoro-β-L-fucose (GDP-CF_3-Fuc)) instead of the methyl to destabilize the oxocarbenium ion-like transition state and decrease the rate of enzyme turnover[39,40], as demonstrated by the reduced Fuc transfer in enzyme assays using GDP-CF_3-Fuc as donor (Extended Data Fig. 6). This derivative has been reported previously[41], but no structures of complexes of this inhibitor with a fucosyltransferase have been published. Using GDP-CF_3-Fuc, we solved the 1.4-Å resolution crystal structure of a FUT9–GDP-CF_3-Fuc–H-type 2 Michaelis complex (Fig. 4a and Supplementary Table 1). The GDP-CF_3-Fuc donor binds to a conserved pocket formed between helices α4, α6 and strands β7, β9 in the donor domain (Fig. 4a,b). With the exception of the diphosphate, the electron density of the GDP-CF_3-Fuc donor was unambiguous (Fig. 4c). The diphosphate electron density reveals a mixture of two conformations, with the best-resolved species representing the intact donor GDP-CF_3-Fuc, and a minor species adopting the diphosphate conformation observed in the GDP-bound complexes FUT9–GDP, FUT9–GDP–H-type 2 and FUT9–GDP–LNnT (Fig. 3a,d). The intact GDP-CF_3-Fuc refined to 75% occupancy, which suggests that the GDP-CF_3-Fuc undergoes a slow hydrolysis in the Michaelis complex. Both the GDP-CF_3-Fuc donor and GDP nucleosides adopt the same conformation with the nucleobase glycosidic torsion angle (O4′–C1′–N9–C4) in the *anti*-rotamer (−94°) conformation and the ribose ring pucker in the favorable $^2T_3$ twist conformation[42,43]. Interactions with the nucleotide base include hydrogen bonds with the backbone amide and carbonyls of Val226, carboxylate of Asp228

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and side chain hydroxyl of Ser196 (Fig. 4c,d). The ribose O4 hydroxyl forms hydrogen bonds with the backbone carbonyl of Val192 and the side chain hydroxyl of Tyr241, while the O3 ribose hydroxyl is hydrogen bonded with Glu255 Oε2. The importance of the Glu255 interaction was tested with the E255A substitution, which reduced k_{cat}/K_m by 288-fold (GDP-Fuc donor kinetics with LNnT as acceptor; Supplementary Table 3).

In the intact FUT9–GDP-CF_3-Fuc–H-type 2 Michaelis complex, the α-phosphate accepts hydrogen bonds from the Ser194 and Asn195 backbone amides and forms a salt bridge with Arg202, while the β-phosphate accepts hydrogen bonds from the Nδ2 atoms of Asn195 and Asn246 (Fig. 4c,d). On hydrolysis, the diphosphate changes conformation by rotating about the nucleotide sugar γ, β, α, ν and μ dihedral angles by −14°, −13°, −63°, 58° and 35°, respectively (Fig. 4e–i). This conformational change rotates the α-phosphate in GDP-CF_3-Fuc away from Arg202 (breaking the salt bridge), and rotates the GDP β-phosphate toward Arg202 (forming a new salt bridge) (Fig. 5a,b,d,e). The rotated β-phosphate also forms an additional salt bridge with the Nζ atom of Lys256. The remaining α- and β-phosphate hydrogen bonding interactions with FUT9 are retained, albeit with different phosphate oxygens (Fig. 5d,e). This change in diphosphate conformation on hydrolysis is probably driven by the relaxation of the strained ν dihedral angle in GDP-CF_3-Fuc (Fig. 4f–i; see catalytic mechanism discussion below). The importance of these diphosphate interactions was examined by R202A and K256A substitutions, which were inactive, and the N195A and N246A substitutions, which reduced k_{cat}/K_m by 82- and 70-fold, respectively (GDP-Fuc donor kinetics with LNnT as acceptor; Supplementary Table 3).

The CF_3-Fuc residue is well ordered and adopts a standard $^1C_4$ chair conformation. The Fuc residue forms hydrogen bonds between the O2 hydroxyl and both Asn246 Nδ2 and Glu137 Oε1 atoms, the O3 hydroxyl and the Tyr168 and Tyr252 hydroxyls, and the O4 hydroxyl and Glu255 Oε1 and Lys256 Nζ atoms (Fig. 4c,d). The C5 trifluoro group is solvent exposed and all three fluorine atoms are well resolved in the electron density (Fig. 4c). The E137A, N246A, T254A and E255A substitutions of fucose-interacting residues exhibited 2,870-, 109-, 394- and 367-fold reductions in k_{cat}/K_m, respectively (kinetics using H-type 2 as acceptor; Supplementary Table 3). The effect of the T254A substitution was rather unexpected for a relatively small change in packing (Fig. 4c,d). Structural analysis reveals that Thr254 also caps the α6 helix, and the mutation probably destabilizes the N terminus of the helix adjacent to the active site.

The conformation of the FUT9 sugar nucleotide donor and the interactions with the donor domain are highly conserved with that of HpFucT (Supplementary Fig. 3h). Most residues interacting with the sugar donor are identical and positioned similarly between FUT9 and HpFucT, including FUT9 side chains Val192, Ser194, Asn195, Arg202, Val226, Asn246, Tyr252, Thr254, Glu255 and Lys256. Differences in donor interactions include Lys225 in HpFucT, which interacts with the O3 hydroxyl of the donor ribose (replaced with Leu231 in FUT9), and Asp228 in FUT9 (missing in HpFucT because of a shorter $Loop_{β8–α5}$). Mutagenesis of FUT9 Asp228 (D228A) led to a 39-fold reduction in k_{cat}/K_m (GDP-Fuc donor kinetics with LNnT as acceptor; Supplementary Table 3). An additional hydrogen bond between FUT9 Tyr168 and the C3 hydroxyl of the donor fucose is replaced by Leu124 in HpFucT, which does not interact with the donor. Notably, the donor fucose in

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

both enzymes sits in the cleft between the donor and acceptor domains and, for the FUT9–GDP-CF_3-Fuc–H-type 2 ternary complex, is sandwiched between the face of the donor-binding domain and the acceptor glycan, with the C6 position extended into the solvent (Fig. 4a–c). This explains the ability of GT10 fucosyltransferases to employ C6-modified GDP-Fuc derivatives as effective donors in glycan-tagging reactions on glycoproteins and cell surfaces[44,45].

Catalytic mechanism

The general mechanism for inverting glycosyltransferases such as FUT9 involves a base-catalyzed deprotonation of the acceptor hydroxyl nucleophile, in concert with an S_N2 attack on the anomeric carbon of the donor to invert the stereochemistry[29,31,35]. The negative charges on the nucleotide diphosphate leaving group are stabilized by positively charged side chains in the active site. In FUT9, the catalytic base Glu137 forms a complex pattern of interactions that deprotonate the GlcNAc O3 nucleophile while also forming a bridging interaction with the donor sugar (Fuc O2) and the acceptor Gal O6 (Fig. 5c). This unique set of bridging interactions provides specificity for the respective donor and acceptor LacNAc substrates by positioning the donor, the acceptor and the catalytic machinery in the active site. The acceptor O3 nucleophile is ideally positioned in close van der Waals contact (3.3 Å) for an S_N2 attack on the anomeric C1 atom of the fucose. The ν dihedral angle of the GDP-CF_3-Fuc diphosphate adopts an angle of 129°, which introduces a steric clash between the O2A α-phosphate oxygen and the larger β-phosphorus atom P_B (Fig. 4f,g). Our analysis reveals a similarly strained ν dihedral angle (138°) in HpFucT (PDB 2NZY). Following hydrolysis, the strained ν dihedral angle in FUT9 relaxes to the favorable staggered conformation (58°) observed in the FUT9 GDP complexes (Fig. 4f,h). In addition to satisfy the same electrostatic interactions with the GDP-CF_3-Fuc donor in the active site, the relaxed GDP conformation forms an additional salt bridge between the β-phosphate and the Nζ atom of Lys256 (Fig. 5d).

Discussion

The high degree of sequence similarity between the mammalian GT10 fucosyltransferases, particularly within the donor-binding domain, suggests that donor interactions are probably identical for all mammalian GT10 isoforms. This hypothesis was supported by a similar positioning for the sugar donor and interacting residues even for the remotely related HpFucT–GDP-Fuc complex (21% sequence identity) (Extended Data Fig. 3h). Distinctions in acceptor specificity among the GT10 fucosyltransferases presumably arise from differences in N-terminal acceptor domain residues facing the substrate binding cleft. While the lack of structural data for an HpFucT–acceptor complex leaves details of acceptor recognition unresolved, the acceptor domain architecture for the bacterial enzyme is completely different from FUT9, except for the position of the catalytic base (Extended Data Fig. 3i).

For human GT10 fucosyltransferases, the lack of a broader collection of enzyme–acceptor complex structures also limits understanding of how structural differences in substrate interactions can drive selective acceptor affinities. For FUT9, the terminal type 2 LacNAc

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

units were identically positioned for H-type 2 and LNnT complexes, while other substrate residues faced the solvent. LacNAc interactions were restricted largely to hydrophobic packing, while Glu137 has a complex pattern of interactions with the donor Fuc residue, the acceptor Gal and as catalytic base interacting with the GlcNAc O3. All these residues, except Phe329, are conserved between the mammalian GT10 enzymes suggesting that these enzymes employ a similar mode of LacNAc interaction and catalytic machinery. However, this raises a question regarding the origin of their differences in acceptor specificity, especially for the bifunctional α1,3/4-fucosyltransferases, FUT3 and FUT5.

FUT9 has one of the simplest specificities among the GT10 enzymes with a preference for type 2 LacNAc or H-type 2 structures, and only trace activity toward sialyl-type 2 (Supplementary Table 3), but not type 1 LacNAc or sialyl-type 1 acceptors[9,46]. FUT3 and FUT5, in contrast, are more promiscuous and can act on type 1 or type 2 LacNAc structures, including H-type and sialyl-type 1 and 2 structures[25,27]. Differences between type 1 and type 2 LacNAc structures include both linkage positions and a 180° flip of the GlcNAc sugar ring relative to the adjoining Gal residue (Fig. 6). Thus, superposition of the Gal residue in type 1 LacNAc with the terminal Gal in the FUT9–GDP–LNnT complex leads to a broadly similar position for the GlcNAc sugar ring, including an identical positioning of the GlcNAc O4 in the type 1 LacNAc relative to the GlcNAc O3 acceptor nucleophile in the type 2 chain complex. Thus, FUT3 and FUT5 probably modify type 1 chains through an analogous set of interactions to those employed in the FUT9–GDP–LNnT complex, but instead present the GlcNAc O4 as the nucleophile (Fig. 6). This suggests that the bifunctional FUT3 and FUT5 active sites would favor GlcNAc ring positions for both type 1 or type 2 chains, while type 1 chain interactions would be disfavored for FUT9. While most donor and acceptor residues are conserved between FUT9, FUT3 and FUT5 on the basis of sequence alignments (Extended Data Fig. 1), three sets of flanking residues are not conserved (Fig. 6). In FUT9, Gln75 extends toward the GlcNAc O6 in the type 2 LacNAc complex, but might cause a steric clash with the GlcNAc *N*-acetyl for a type 1 chain complex (Fig. 6). The equivalent residues in FUT3 and FUT5 are the shorter side chains of Ile74 or Thr87, respectively (only the FUT3 structure is shown in Fig. 6). Another active site difference is His141 in FUT9 in proximity to the acceptor GlcNAc *N*-acetyl, which is replaced by Asn142 or Asn155 in FUT3 and FUT5, respectively. These latter residues may more effectively accommodate the shorter GlcNAc O6 hydroxyl for the type 1 chain acceptor (Fig. 6). Finally, the side chain of Arg111 extends toward the subterminal Gal residue in the type 2 chain complex, but the equivalent residues in FUT3 and FUT5 are Trp111 and Trp124, respectively. This latter sequence difference was noted as a part of a VxxHH(W/R)(D/E) acceptor-binding motif that correlated broadly with either bifunctional (Trp) or type 2 chain (Arg) fucosyltransferase activity[47]. Mutagenesis of FUT5 and FUT6 to swap these residues did not exchange acceptor specificities, suggesting additional residues were involved in the substrate selectivity[47]. Similarly, our mutagenesis data with a FUT9 R111W mutation resulted in an 89-fold reduction in activity toward H-type 2 and only trace levels of activity toward H-type 1 structures (Supplementary Table 3), while a combined R111W/H141N mutation resulted in a 56-fold reduction in activity toward H-type 2 and no activity toward H-type 1. Thus, additional residues within the acceptor-binding domain are clearly required to support selective acceptor recognition that either enhances or blocks

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

interactions with modified type 1 and type 2 LacNAc structures within the framework of an adjacent, highly conserved donor-binding domain.

By comparison, equivalent modeling of type 1 or type 2 chains in the active site of HpFucT indicates that the bacterial enzyme has a broad acceptor-binding cleft and only one residue beyond the catalytic base (Glu137 in FUT9) that is positioned to interact with a similarly positioned acceptor structure (Fig. 6). These data suggest either a conformational change in the enzyme active site or that alternative substrate interactions are employed by HpFucT.

GT10 fucosyltransferases are present in bacteria, eukaryotes and viruses, and all family members belong to a broad subcluster of GT-B fold enzymes (GT-B0 cluster) based on a recent convolutional neural network analysis[48]. This cluster consists of at least 16 other CAZy families with diverse specificities for donor and acceptor recognition and catalytic mechanism, and some members of this cluster were identified in the Dali structural similarity analysis (Supplementary Table 2). Enzymes within this larger gene cluster are predominantly conserved only in their C-terminal donor-binding domains, but contain highly divergent N-terminal sequences involved in diverse acceptor interactions. As a result, mammalian GT10 fucosyltransferase donor and acceptor specificity can be seen within a broader continuum of GT-B fold glycosyltransferase evolution. Mammalian GT10 enzymes employ nearly identical modes of sugar donor interactions paired with subtle differences in acceptor domain structures that drive distinct acceptor selectivities for Lewis antigen synthesis. Bacterial GT10 enzymes, such as HpFucT, evolved more divergent acceptor-binding domain structures while retaining the general architecture and interacting residues of the donor-binding domain to synthesize mimics of host Lewis antigens on their lipopolysaccharides. Further diversity of enzymes within the broader GT-B0 cluster retained the general structural elements of the donor-binding domain, but evolved alternative sugar donor specificities and paired those with flanking N-terminal sequences harboring highly divergent acceptor specificities. Thus, the modular and independent evolution of donor- and acceptor-binding domains among the GT-B enzymes has allowed the rapid diversification of enzyme specificities involved in synthesis of complex glycan structures in biological systems.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41589-023-01345-y.

Methods

Expression construct design and purification of human FUT9

A fusion protein expression construct encoding the catalytic domain of human FUT9 (Uniprot Q9Y231, residues 39–359) was generated by amplifying a FUT9 Mammalian Gene Collection clone and inserting it into the pDONR221 vector using BP Clonase II Gateway recombination[32]. The amplification product in the pDONR221 vector was transferred to

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

both the pGEn2-DEST and pGEn3-DEST mammalian expression vectors by LR Clonase Gateway recombination to form FUT9-pGEn2 that encodes an N-terminal signal sequence followed by an 8xHis tag, AviTag recognition site, 'superfolder' GFP, the TEV protease recognition site and the coding region of human FUT9 downstream of the cytomegalovirus promoter[32]. The FUT9-pGEn3 construct contains an Fc domain of human immunoglobulin G (IgG) between the 'superfolder' GFP domain and the TEV protease recognition site, and provided improved binding to the Ni^{2+}-NTA column during purification. Both the proteins were expressed by transient transfection in wild-type HEK293F cells (Freestyle 293F, ThermoFisher Scientific) for kinetic studies (FUT9-pGEN2), or mutant HEK293S (GnTl-) cells[33] (ATCC, catalog no. CRL-3022) for structural studies (FUT9-pGEn2 and FUT9-pGEn3). The procedure for transient transfection was performed as reported previously[33]. Briefly, cells were suspended in Free-style 293 Expression Medium (Life Technologies) and Ex-Cell 293 Serum-free medium (Sigma) at a ratio of 9:1 (9:1 media) and a cell density of 3–3.5 × 10^6 cells ml^{-1}. Transfection was performed using polyethyleneimine (linear 25 kDa PEI, Polysciences) as the transfection agent at a concentration of 5 μg ml^{-1} and 4 μg ml^{-1} of expression vector DNA and incubated at 37 °C in a humidified CO_2 shaker incubator (150 r.p.m.) for 24 h. The cell suspension was then diluted with an equal volume of 9:1 media containing valproic acid (VPA, Sigma) to a final concentration of 2.2 mM and cultured for an additional 5 d at 37 °C before collection. Kinetic studies were performed on active site mutants generated using the Q5 site-directed mutagenesis kit (New England Biolabs) and mutations were confirmed by DNA sequencing. For kinetic analysis, both wild-type and mutant forms of FUT9 were expressed in the pGEn2 vector as secreted soluble fusion proteins.

The collected supernatants were passed through a 0.8-μm membrane filter and loaded onto a Ni^{2+}-NTA column (Qiagen) pre-equilibrated with 20 mM HEPES, pH 7.4, 300 mM NaCl, 20 mM imidazole and 5% glycerol. The column was washed with the same buffer containing 50 mM imidazole to remove nonspecific protein interactions and the FUT9-GFP fusion protein was eluted using 300 mM imidazole. For enzyme kinetics, the protein was pooled and concentrated following the Ni^{2+}-NTA column chromatography and assays were performed. For structural studies the FUT9-pGEn2 or FUT9-pGEn3 constructs were expressed in HEK293S (GnTl-) cells, purified by Ni^{2+}-NTA chromatography and the GFP-Fc tag and N-glycans were cleaved by treating the concentrated protein for 16 h at 4 °C with recombinant TEV protease and EndoF1 as previously described[33] at a 1:5 ratio relative to the recombinant FUT9 fusion protein. The sample was then loaded onto a Ni^{2+}-NTA column to remove the GFP and Ig Fc fusion tags and the His-tagged TEV and EndoF1 enzymes. The resulting untagged FUT9 in the flow-through fraction was further purified using a Superdex 75 column (GE Healthcare) pre-equilibrated with 20 mM HEPES, pH 7.4, 150 mM NaCl and 60 mM imidazole, and peak fractions were collected.

Size-exclusion chromatography–multiangle scattering analysis

The oligomerization of the TEV protease and EndoF cleaved FUT9 preparation was analyzed on an analytical grade Superdex 75 column pre-equilibrated with 25 mM HEPES, pH 7.4, 150 mM NaCl and 0.02% NaN_3 by injecting 20 μl of a 1 mg ml^{-1} enzyme preparation. Light scattering was detected using a miniDAWN TREOS detector (Wyatt

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Technology), and differential refractive index was measured using an Optilab rEX detector (Wyatt Technology). The data were further analyzed using the ASTRA v.6.0 software (Wyatt Technology).

FUT9 kinetic analysis

Enzyme assays for FUT9 were performed using calf intestinal alkaline phosphatase (CIAP, Promega)-treated GDP-Fucose (Promega) as donor[17] at a final concentration of 0.2 mM, and kinetics for the wild-type and mutant forms of FUT9 were determined using the GDP-Glo Glycosyltransferase assay (Promega) as described previously[17]. Briefly, reactions were performed in a 10 μl reaction volume consisting of 0.2 mM GDP-Fucose as donor and an acceptor substrate (LNnT (Biosynth Carbosynth), H-type 2 trisaccharide (Biosynth Carbosynth), type 2 LacNAc-O-Me or 3′-α-sialyl-type 2 LacNAc-O-Me, type 1 LacNAc-O-Me or 3′-α-sialyl-type 1 LacNAc-O-Me synthesized as described in Supplementary Information) at varied concentrations from 0 to 1 mM. Assays were performed at 37 °C for 30 min in a universal buffer (200 mM each of Tris, MES, MOPS, pH 7.5) using either wild-type or mutant FUT9-GFP fusion proteins. Donor kinetics were examined using GDP-Fucose (0–0.2 mM) with 1 mM LNnT as acceptor. Reactions were stopped using 5 μl of GDP detection reagent at a 1:1 ratio in white 384-well plates (Corning) and incubated in the dark for 1 h at room temperature. The luminescence values were detected using a GloMax-Multi detection plate reader (Promega) and nonlinear curve fitting and steady-state parameters of K_m and k_{cat} values were analyzed using a GDP standard curve in GraphPad Prism 6 software. All assays were performed at least in duplicate.

Crystallization and X-ray data collection

Crystals of FUT9 were grown from purified enzyme preparations and premixed with GDP-Fucose (0.1 mM), GDP (3 mM), LNnT (10 mM), H-type 2 (10 mM) or GDP with acceptor. The heavy atom derivative was prepared using FUT9 crystals grown in the presence of 0.1 mM GDP-Fucose by soaking the crystals in a 0.6 M solution of CsCl and NaI (1:1 ratio). The FUT9–GDP-CF3-Fuc–H-type 2 and FUT9–GDP–H-type 2 complexes were obtained by soaking the ligands (33–66 mM GDP-CF3-Fuc and 12–20 mM H-type 2) in ligand-free FUT9 crystals for 4 and 21 h, respectively. X-ray diffraction data were collected at the SER-CAT 22-ID beamline, Argonne National Laboratory, with the Rayonix 300HS or Eiger 16M detectors using wavelengths of 1.7 Å (FUT9–HA) or 1.0 Å (native, ligand-free FUT9 and its complexes) at a temperature of 100 K. Data were processed using XDS[51].

Structure solution and refinement

Data from the heavy atom-derivatized (Cs and I) FUT9 were phased using single-wavelength anomalous diffraction. The calculated protein phases for nine cesium and fourteen iodide sites gave a figure of merit of 0.73. Iterative cycles of automated model building and density modification produced the initial model for FUT9–HA. The crystal structures of ligand-free FUT9 and its complexes were solved by molecular replacement using Phaser[52] and FUT9–HA as the search model (Supplementary Table 1). Structure refinement using Phenix[52] and iterative manual fitting using Coot[53] produced all of the final models, which are devoid of Ramachandran outliers (Supplementary Table 1). Disordered

residues in the NH_2 terminus (residues 38–62) were not modeled and *B*-factors were refined using either TLS or anisotropic refinement depending on the data resolution.

Extended Data

a

b

	HpFucT (2NZY)	FUT9	FUT4	FUT7	FUT6	FUT5
FUT9	21%					
FUT4	19%	47%				
FUT7	21%	42%	44%			
FUT6	18%	42%	45%	45%		
FUT5	19%	43%	45%	44%	84%	
FUT3	19%	43%	43%	45%	79%	85%

Extended Data Fig. 1 |. Structural alignment of GT10 fucosyltransferases using the PROMALS3D server.

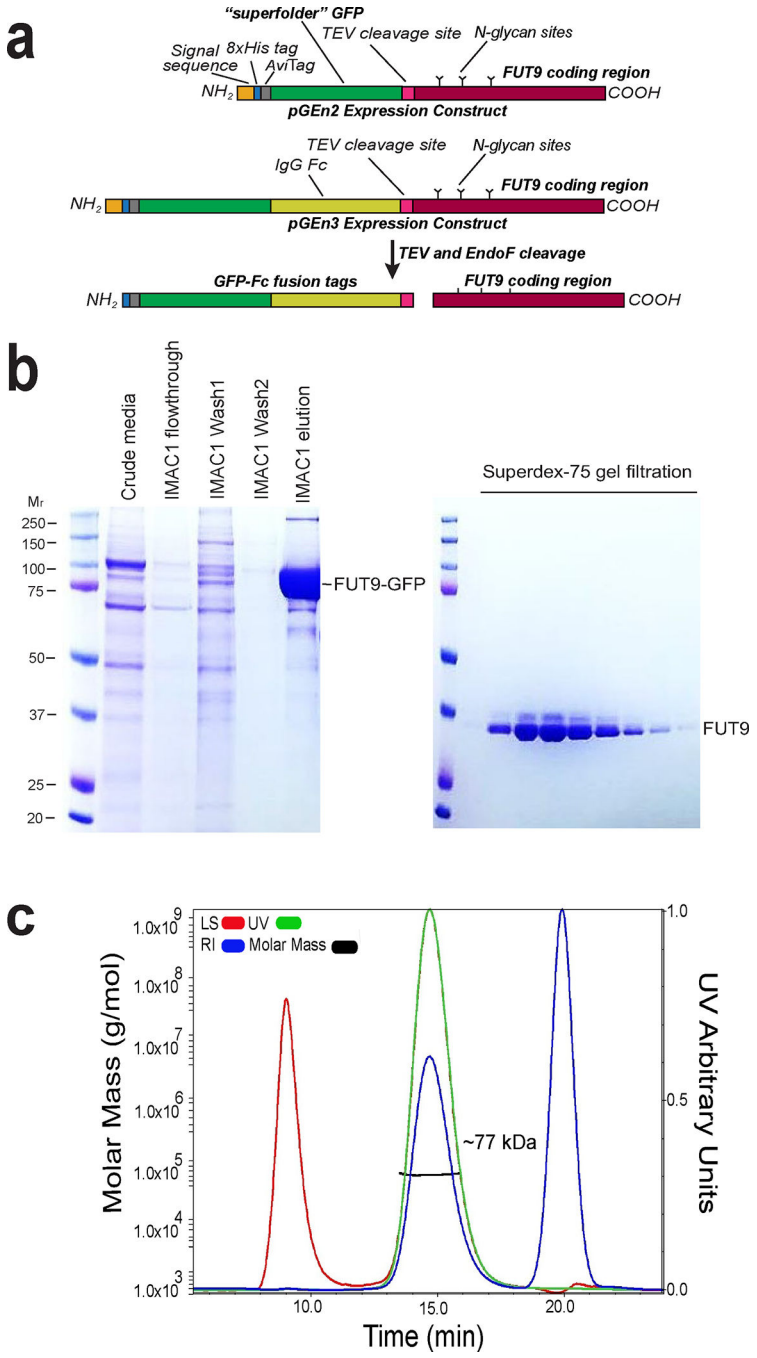
(***a***) The structure of the human FUT9:GDP:LNnT complex was aligned with the HpFucT (PDB 2NZY[29]) and the primary sequences of FUT3 (UniProt P21217), FUT4 (UniProt

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

P22083), FUT5 (UniProt Q11128), FUT6 (UniProt P51993), and FUT7 (UniProt Q11130) using PROMALS 3D[54]. Conserved secondary structure elements are indicated as helices (*h*) or beta strands (*e*) and consensus amino acid positions are classified by amino acid character (aliphatic (l), aromatic (@), hydrophobic (h), alcohol (o), polar (p), tiny (t), small (s), bulky (b), positively charged (+), negatively charged (−), and charged (c)) or as bold uppercase for conserved. Red boxes represent the residues within the acceptor binding domain of FUT9. Blue boxes represent residues in the donor binding domain of FUT9. Residues involved in acceptor interactions for FUT9 are indicated with yellow stars. Residues involved in donor interactions for FUT9 are indicated by red stars. The catalytic base (E137) is indicated with a green arrow. Putative residues that contribute to discrimination between type 2 and type 1 LacNAc unit recognition are indicated with red circles. The conserved N-glycan site is indicated by a black box. Cys residues forming disulfide bonds in FUT9 (and conserved Cys residues forming disulfides in other human GT10 fucosyltransferases) are indicated by green boxes. (***b***) Percent identity between the respective GT10 fucosyltransferases based on structure and sequence alignments using the Dali server[36].

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



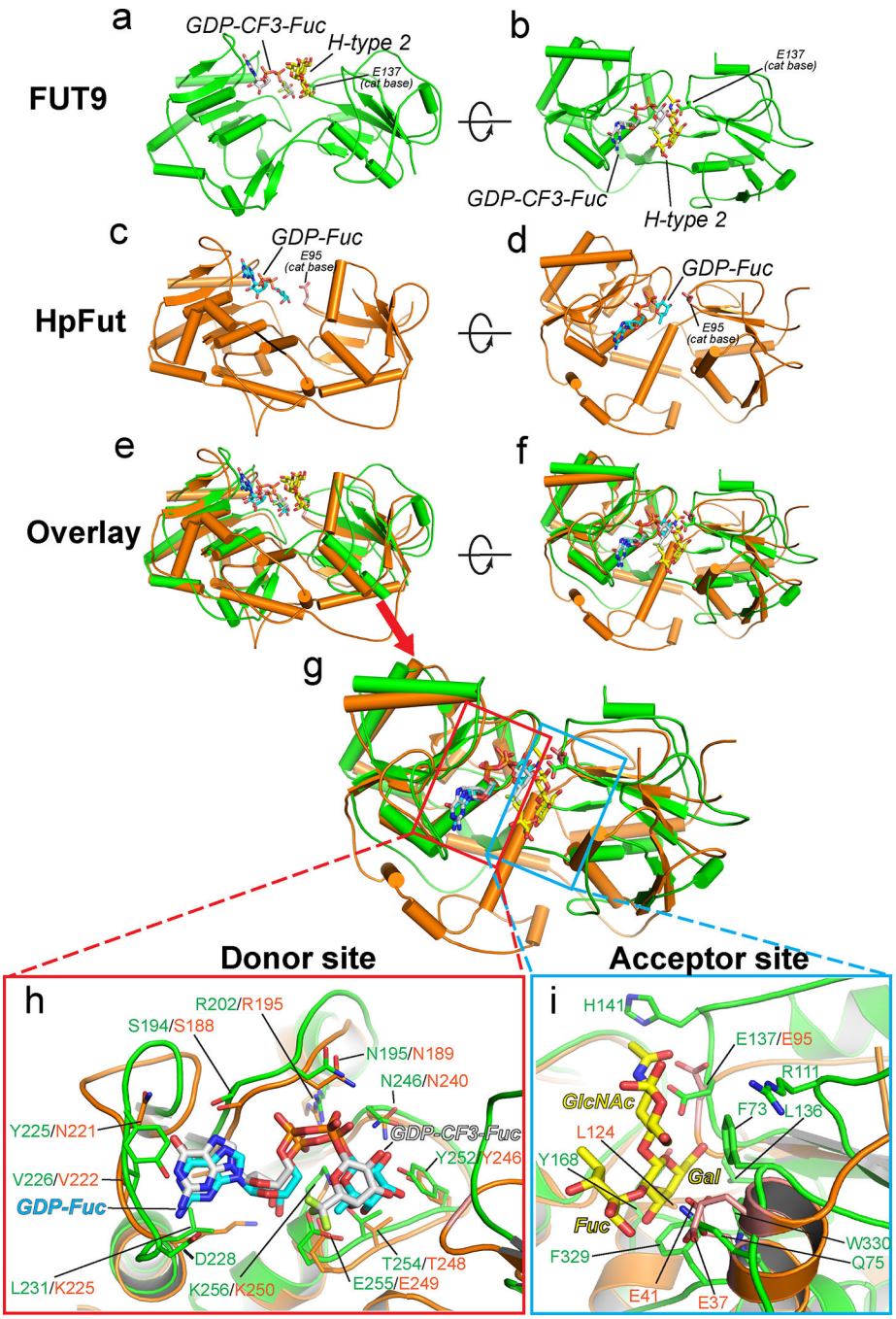
Extended Data Fig. 2 |. Expression, structure, and dimerization of the catalytic domain of human FUT9.

(***a***) Diagrammatic representation of the protein coding regions from the FUT9-pGEn2 and FUT9-pGEn3 expression constructs. The fusion protein from the FUT9-pGEn2 construct contains an NH_2-terminal signal sequence followed by an 8xHis tag, AviTag, superfolder GFP, TEV protease cleavage site, and the catalytic domain of FUT9 containing three *N*-glycan consensus sequons sites at N62, N101, and N153. The protein coding region for the FUT9-pGEn3 construct is identical to the FUT9-pGEN2 construct with the exception

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

of the inclusion of an Ig Fc domain between the GFP domain and TEV protease cleavage site. (***b***) Expression of the recombinant product in HEK293S (GnTI-) cells resulted in secretion of the fusion protein into the culture medium (Crude media), and subsequent Ni^{2+}-NTA purification yielded a highly-enriched enzyme preparation (IMAC1 elution). Cleavage of the enzyme with TEV protease and EndoF1 resulted in removal of the tag sequences and glycans, and further purification by Ni^{2+}-NTA chromatography and Superdex-75 (Superdex-75 gel filtration) led to the final purified preparation. Each transfection experiment was performed at least three times and two different SDS PAGE gels of the samples were generated. Data presented are representative of the respective experiments Original uncropped images are provided in the **Source Data**. (***c***) The purified FUT9 catalytic domain was further characterized by size exclusion-multiangle light scattering (SEC-MALS). A_{280} is shown by the green line, refractive index in blue, light scattering in red and calculated molar mass in black. The molecular mass derived from SEC-MALS analysis (~77 kDa) is in close agreement with a dimeric form of the FUT9 catalytic domain monomer following cleavage with TEV and EndoF1 (~37.8 kDa).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

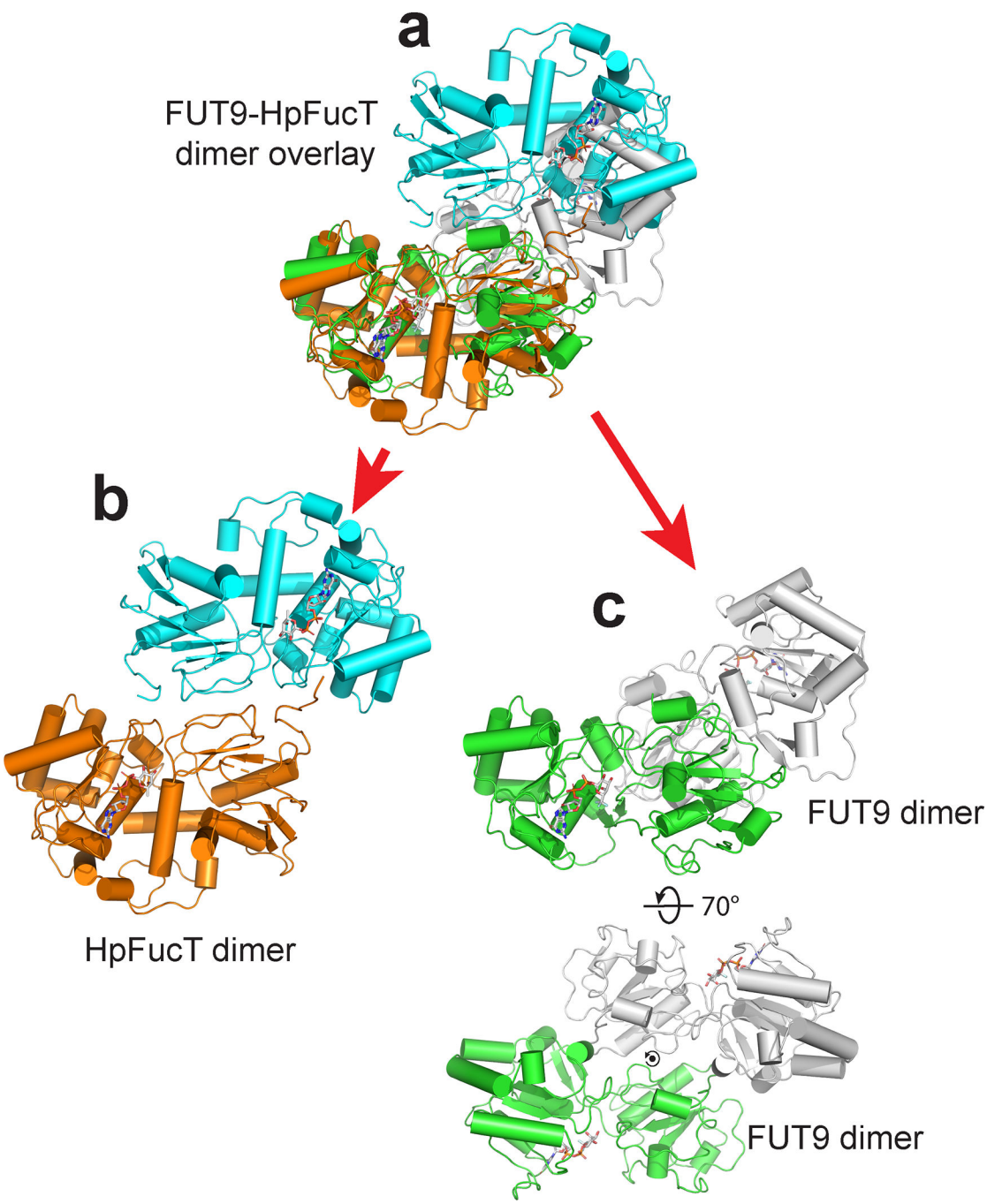


Extended Data Fig. 3 |. Structural alignment of human FUT9 with *H. pylori* FucT.

(*a, b*) Structure of FUT9:GDP-CF_3-Fuc:H-type 2 (green cartoon for protein, white and yellow sticks for the donor and H-type 2, respectively) and ***(c, d)*** HpFucT:GDP-Fuc complex (PDB 2NZY[29], orange cartoon for protein and cyan sticks for GDP-Fuc) were (***e, f***) aligned based on structural similarity of the donor binding domain. **(*g*)** A zoom in view of the aligned structures are highlighted for the donor binding site (red box) and the acceptor binding site (blue box). (***h***) The donor binding site is further displayed with interacting residues shown as thin sticks (FUT9 as green sticks and HpFucT as orange sticks) and the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

GDP-CF3-Fuc as white sticks and GDP-Fuc as cyan sticks). (***i***) The acceptor binding site is further displayed with interacting residues shown as thin sticks colored as in panel ***h***. The H-type 2 acceptor structure is shown as yellow sticks.

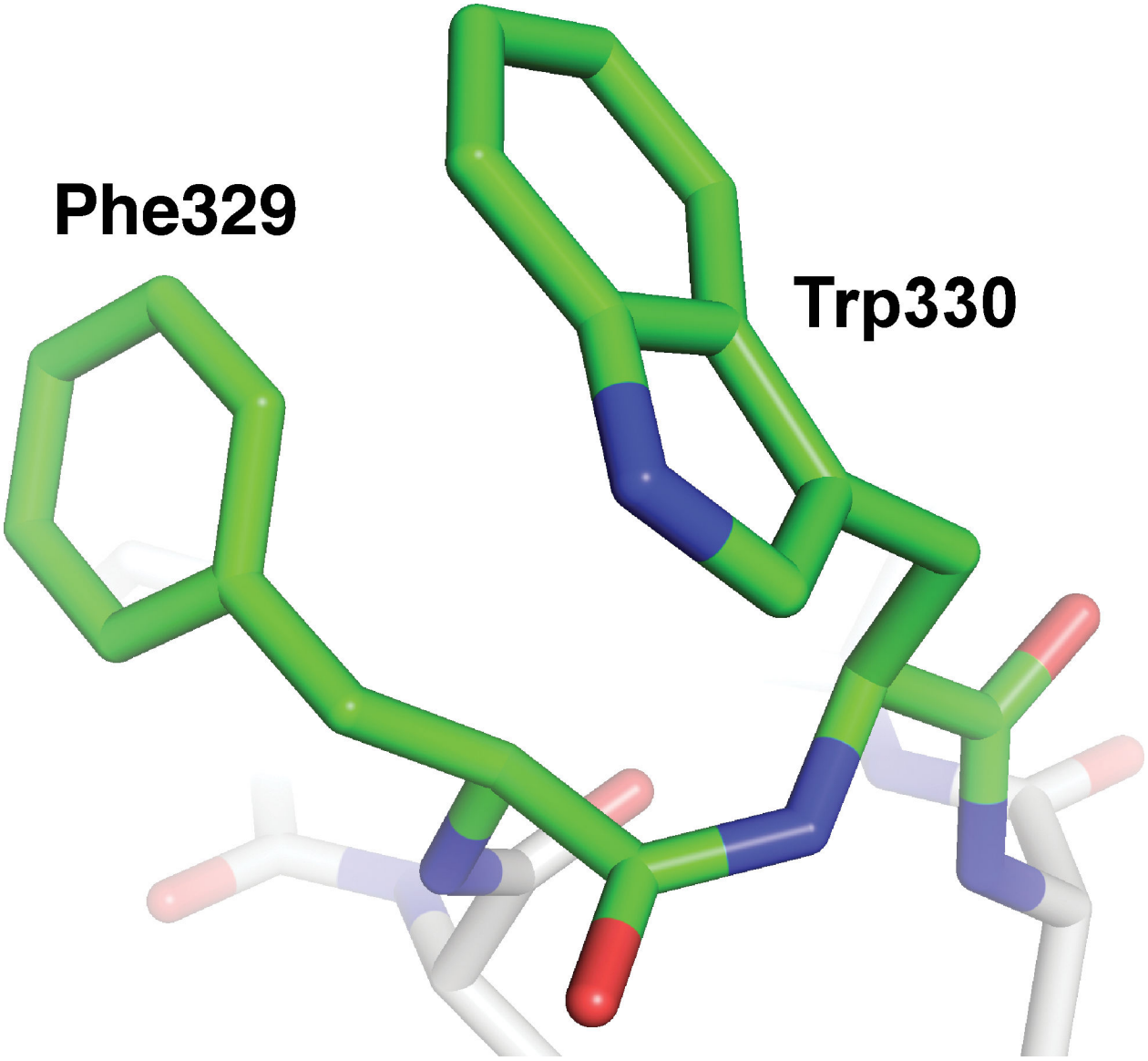


Extended Data Fig. 4 |. Alignment of the dimeric structures of FUT9 and HpFucT.

(*a*) The homodimeric structures of FUT9 (green and gray cartoons for chains A and B, respectively) and HpFucT (orange and cyan cartoons for chains A and B, respectively) were aligned based on the conserved donor domain structures for the respective chain A of each

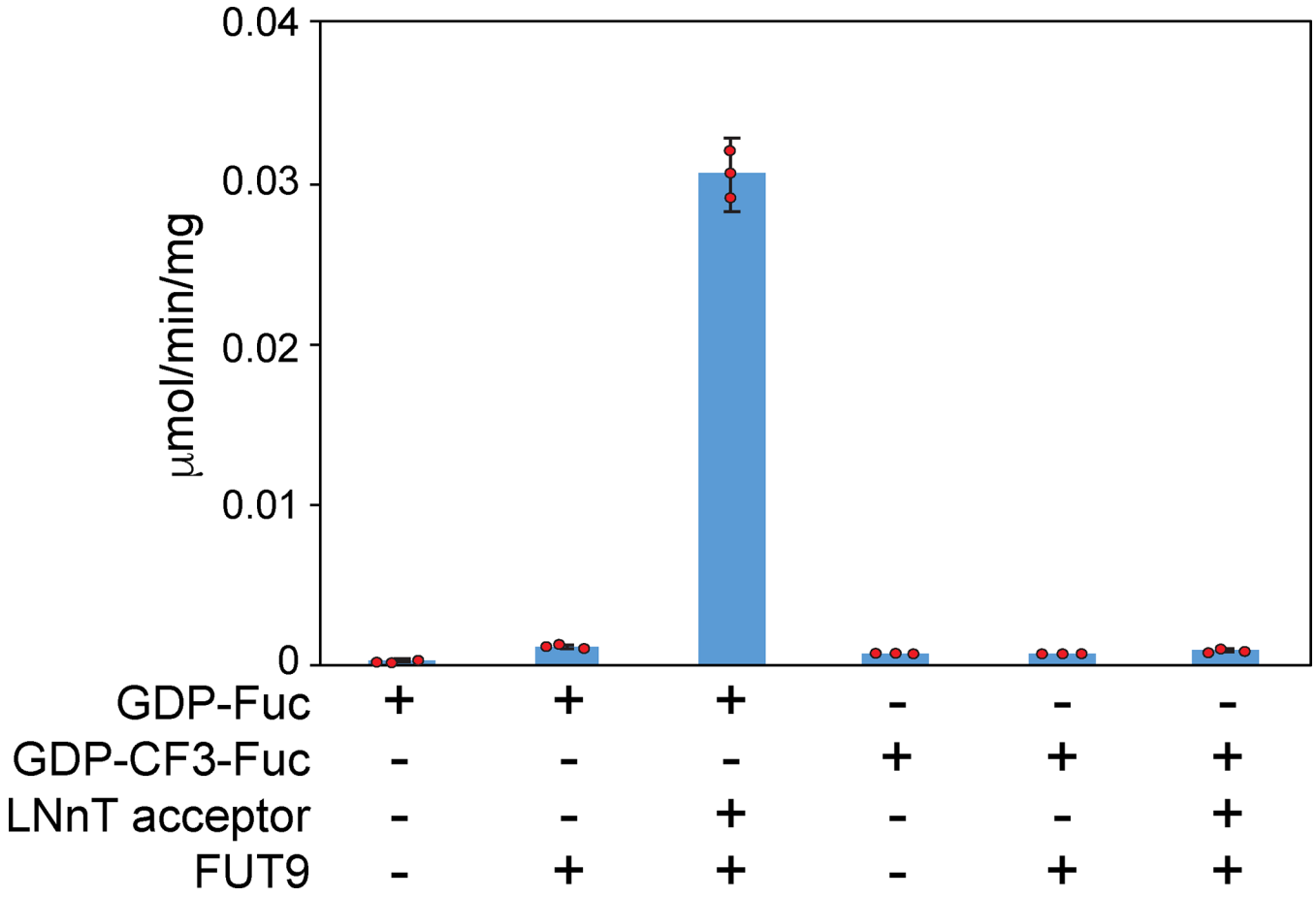
Author Manuscript Author Manuscript Author Manuscript Author Manuscript

homodimer. The respective individual homodimeric structures of (***b***) HpFucT and (***c***) FUT9 (two views rotated by 70°) are shown with the two-fold axes as indicated. The two-fold axis of FUT9 requires a 70° rotation relative to the HpFucT in order to display the 2-fold axis indicating a distinct interface region for the monomers in the respective homodimers.



Extended Data Fig. 5 |. *Cis*-Phe329 in FUT9.
Residue Phe329 which precedes helix α12 is involved in a *cis*-bond (ω angle of ~13°). Stabilizing interactions for the unfavorable bond angle include aromatic and aromatic edge−to-face stacking of Phe329 and Trp330. The *cis*-Phe is not observed in HpFucT due to the lack of structural conservation in the region.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Extended Data Fig. 6 |. Enzyme activity of FUT9 employing GDP-Fuc or GDP-CF3-Fuc donors and LNnT as acceptor.

FUT9 enzyme activity was determined as indicated in the presence or absence of GDP-Fuc or GDP-CF3-Fuc as donors and in the presence or absence of LNnT as acceptor. The reaction was monitored using the GDP-Glo assay format. The increase in signal over the no-enzyme blank reflects donor hydrolysis (in the absence of acceptor) or sugar transfer (in the presence of LNnT as acceptor). Plots show the mean values (bar) +/− s.d. (error bars) for n = 3 technical replicates (red circles).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The work was supported by NIH grants R01GM130915 (K.W.M., Z.A.W.), P41GM103390 (K.W.M., G.-J.B.) and P01GM107012 (K.W.M., G.-J.B.). C.H. was partially supported by NIGMS training grant, T32 GM1007004. S.G.W. was supported by funding from the Canadian Institutes of Health Research (CIHR). We thank the staff at the SouthEast Regional Collaborative Access Team (SER-CAT), which is supported in part by the National Institutes of Health 623 (S10 RR25528 and S10 RR028976). We also acknowledge the National Institute of General Medical Sciences for an equipment grant supporting our in-house X-ray facility (S10 OD021762). We also thank Amgen Inc. for their donation of trifluoromethyl fucose from which we synthesized GDP-CF_3-Fuc. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Data availability

Atomic coordinates and structure factors for the FUT9 structures listed in Supplementary Table 1 have been deposited in the Protein Data Bank as PDB entries 8D0O, 8D0P, 8D0Q, 8D0R, 8D0S, 8D0U, 8D0W and 8D0X. Construct designs, annotations and sequences are summarized on our website (glycoenzymes.ccrc.uga.edu) and plasmids encoding wild-type FUT9 are available from DNASU (dnasu.org). Expression constructs encoding FUT9 mutants are available from the corresponding author on reasonable request. All data generated or analyzed during this study are included in this published article (and its Supplementary Information files) and are available from the corresponding author on reasonable request[54]. Source data are provided with this paper.

References

1. Moremen KW, Tiemeyer M & Nairn AV Vertebrate protein glycosylation: diversity, synthesis and function. Nat. Rev. Mol. Cell Biol. 13, 448–462 (2012). [PubMed: 22722607]
2. Varki A Biological roles of glycans. Glycobiology 27, 3–49 (2017). [PubMed: 27558841]
3. Cummings RD & Pierce JM The challenge and promise of glycomics. Chem. Biol. 21, 1–15 (2014). [PubMed: 24439204]
4. Schneider M, Al-Shareffi E & Haltiwanger RS Biological functions of fucose in mammals. Glycobiology 27, 601–618 (2017). [PubMed: 28430973]
5. Ma B, Simala-Grant JL & Taylor DE Fucosylation in prokaryotes and eukaryotes. Glycobiology 16, 158R–184R (2006).
6. Liang JX, Liang Y & Gao W Clinicopathological and prognostic significance of sialyl Lewis X overexpression in patients with cancer: a meta-analysis. Onco Targets Ther. 9, 3113–3125 (2016). [PubMed: 27307752]
7. Ohmori K et al. P- and E-selectins recognize sialyl 6-sulfo Lewis X, the recently identified L-selectin ligand. Biochem. Biophys. Res. Commun. 278, 90–96 (2000). [PubMed: 11071860]
8. Sperandio M Selectins and glycosyltransferases in leukocyte rolling in vivo. FEBS J. 273, 4377–4389 (2006). [PubMed: 16956372]
9. de Vries T, Knegtel RM, Holmes EH & Macher BA Fucosyltransferases: structure/function studies. Glycobiology 11, 119R–128R (2001).
10. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM & Henrissat B The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42, D490–D495 (2014). [PubMed: 24270786]
11. Breton C, Oriol R & Imberty A Conserved structural features in eukaryotic and prokaryotic fucosyltransferases. Glycobiology 8, 87–94 (1998). [PubMed: 9451017]
12. Costache M et al. Evolution of fucosyltransferase genes in vertebrates. J. Biol. Chem. 272, 29721–29728 (1997). [PubMed: 9368041]
13. Oriol R, Mollicone R, Cailleau A, Balanzino L & Breton C Divergent evolution of fucosyltransferase genes from vertebrates, invertebrates, and bacteria. Glycobiology 9, 323–334 (1999). [PubMed: 10089206]
14. Lowe JB The blood group-specific human glycosyltransferases. Bailliere's Clin. Haematol. 6, 465–492 (1993). [PubMed: 8043935]
15. Scharberg EA, Olsen C & Bugert P The H blood group system. Immunohematology 32, 112–118 (2016). [PubMed: 27834485]
16. Stowell CP & Stowell SR Biologic roles of the ABH and Lewis histo-blood group antigens part I: infection and immunity. Vox Sang. 114, 426–442 (2019). [PubMed: 31070258]
17. Boruah BM et al. Characterizing human alpha-1,6-fucosyltransferase (FUT8) substrate specificity and structural similarities with related fucosyltransferases. J. Biol. Chem. 295, 17027–17045 (2020). [PubMed: 33004438]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

18. Ihara H et al. in Handbook of Glycosyltransferases and Related Genes (eds Taniguchi N et al.) 581–596 (Springer, 2014).
19. Lira-Navarrete E & Hurtado-Guerrero R A perspective on structural and mechanistic aspects of protein O-fucosylation. Acta Crystallogr. F Struct. Biol. Commun. 74, 443–450 (2018). [PubMed: 30084393]
20. Urbanowicz BR et al. Structural, mutagenic and in silico studies of xyloglucan fucosylation in *Arabidopsis thaliana* suggest a water-mediated mechanism. Plant J. 91, 931–949 (2017). [PubMed: 28670741]
21. Van Der Wel H, Fisher SZ & West CM A bifunctional diglycosyltransferase forms the Fucα1,2Galβ1,3-disaccharide on Skp1 in the cytoplasm of *Dictyostelium*. J. Biol. Chem. 277, 46527–46534 (2002). [PubMed: 12244067]
22. Kudo T & Narimatsu H in Handbook of Glycosyltransferases and Related Genes (eds Taniguchi N et al.) 541–547 (Springer, 2014).
23. Kudo T & Narimatsu H in Handbook of Glycosyltransferases and Related Genes (eds Taniguchi N et al.) 573–580 (Springer, 2014).
24. Kudo T & Narimatsu H in Handbook of Glycosyltransferases and Related Genes (eds Taniguchi N et al.) 597–603 (Springer, 2014).
25. Kannagi R in Handbook of Glycosyltransferases and Related Genes (eds Taniguchi N et al.) 549–558 (Springer, 2014).
26. Kannagi R in Handbook of Glycosyltransferases and Related Genes (eds Taniguchi N et al.) 559–571 (Springer, 2014).
27. Kudo T & Narimatsu H in Handbook of Glycosyltransferases and Related Genes (eds Taniguchi N et al.) 531–539 (Springer, 2014).
28. Mollicone R & Oriol R in Handbook of Glycosyltransferases and Related Genes (eds Taniguchi N et al.) 605–622 (Springer, 2014).
29. Sun HY et al. Structure and mechanism of *Helicobacter pylori* fucosyltransferase. A basis for lipopolysaccharide variation and inhibitor design. J. Biol. Chem. 282, 9973–9982 (2007). [PubMed: 17251184]
30. Tan Y et al. Directed evolution of an alpha1,3-fucosyltransferase using a single-cell ultrahigh-throughput screening method. Sci. Adv. 5, eaaw8451 (2019). [PubMed: 31633018]
31. Lairson LL, Henrissat B, Davies GJ & Withers SG Glycosyltransferases: structures, functions, and mechanisms. Annu. Rev. Biochem. 77, 521–555 (2008). [PubMed: 18518825]
32. Moremen KW et al. Expression system for structural and functional studies of human glycosylation enzymes. Nat. Chem. Biol. 14, 156–162 (2018). [PubMed: 29251719]
33. Meng L et al. Enzymatic basis for N-glycan sialylation: structure of rat α2,6-sialyltransferase (ST6GAL1) reveals conserved and unique features for glycan sialylation. J. Biol. Chem. 288, 34680–34698 (2013). [PubMed: 24155237]
34. Krissinel E & Henrick K Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797 (2007). [PubMed: 17681537]
35. Moremen KW & Haltiwanger RS Emerging structural insights into glycosyltransferase-mediated synthesis of glycans. Nat. Chem. Biol. 15, 853–864 (2019). [PubMed: 31427814]
36. Holm L & Rosenstrom P Dali server: conservation mapping in 3D. Nucleic Acids Res. 38, W545–W549 (2010). [PubMed: 20457744]
37. Ve T, Williams SJ & Kobe B Structure and function of Toll/interleukin-1 receptor/resistance protein (TIR) domains. Apoptosis 20, 250–261 (2015). [PubMed: 25451009]
38. Martinez-Duncker I, Mollicone R, Candelier JJ, Breton C & Oriol R A new superfamily of protein-*O*-fucosyltransferases, α2-fucosyltransferases, and α6-fucosyltransferases: phylogeny and identification of conserved peptide motifs. Glycobiology 13, 1C–5C (2003). [PubMed: 12634318]
39. Burkart MD et al. Chemo-enzymatic synthesis of fluorinated sugar nucleotide: useful mechanistic probes for glycosyltransferases. Bioorg. Med. Chem. 8, 1937–1946 (2000). [PubMed: 11003139]
40. Dai YW et al. Synthetic fluorinated l-fucose analogs inhibit proliferation of cancer cells and primary endothelial cells. ACS Chem. Biol. 15, 2662–2672 (2020). [PubMed: 32930566]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

41. Allen JG et al. Facile modulation of antibody fucosylation with small molecule fucostatin inhibitors and cocrystal structure with GDP-mannose 4,6-dehydratase. ACS Chem. Biol. 11, 2734–2743 (2016). [PubMed: 27434622]
42. Agirre J Strategies for carbohydrate model building, refinement and validation. Acta Crystallogr. D Struct. Biol. 73, 171–186 (2017). [PubMed: 28177313]
43. Cremer D & Pople JA General definition of ring puckering coordinates. J. Am. Chem. Soc. 97, 1354–1358 (1975).
44. Hong S et al. Direct visualization of live zebrafish glycans via single-step metabolic labeling with fluorophore-tagged nucleotide sugars. Angew. Chem. Int. Ed. Engl. 58, 14327–14333 (2019). [PubMed: 31295389]
45. Liu Z et al. Detecting tumor antigen-specific T cells via interaction-dependent fucosyl-biotinylation. Cell 183, 1117–1133.e19 (2020). [PubMed: 33096019]
46. Dupuy F et al. A single amino acid in the hypervariable stem domain of vertebrate α1,3/1,4-fucosyltransferases determines the type 1/type 2 transfer. Characterization of acceptor substrate specificity of the Lewis enzyme by site-directed mutagenesis. J. Biol. Chem. 274, 12257–12262 (1999). [PubMed: 10212193]
47. Dupuy F, Germot A, Julien R & Maftah A Structure/function study of Lewis α3- and α3/4-fucosyltransferases: the α1,4 fucosylation requires an aromatic residue in the acceptor-binding domain. Glycobiology 14, 347–356 (2004). [PubMed: 14718375]
48. Taujale R et al. Mapping the glycosyltransferase fold landscape using interpretable deep learning. Nat. Commun. 12, 5656 (2021). [PubMed: 34580305]
49. Jumper J et al. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589 (2021). [PubMed: 34265844]
50. Kirschner KN et al. GLYCAM06: a generalizable biomolecular force field. Carbohydrates. J. Comput. Chem. 29, 622–655 (2008). [PubMed: 17849372]
51. Kabsch W XDS. Acta Crystallogr. D Biol. Crystallogr. 66, 125–132 (2010). [PubMed: 20124692]
52. Adams PD et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221 (2010). [PubMed: 20124702]
53. Emsley P, Lohkamp B, Scott WG & Cowtan K Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501 (2010). [PubMed: 20383002]
54. Pei J, Kim BH & Grishin NV PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic Acids Res. 36, 2295–2300 (2008). [PubMed: 18287115]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

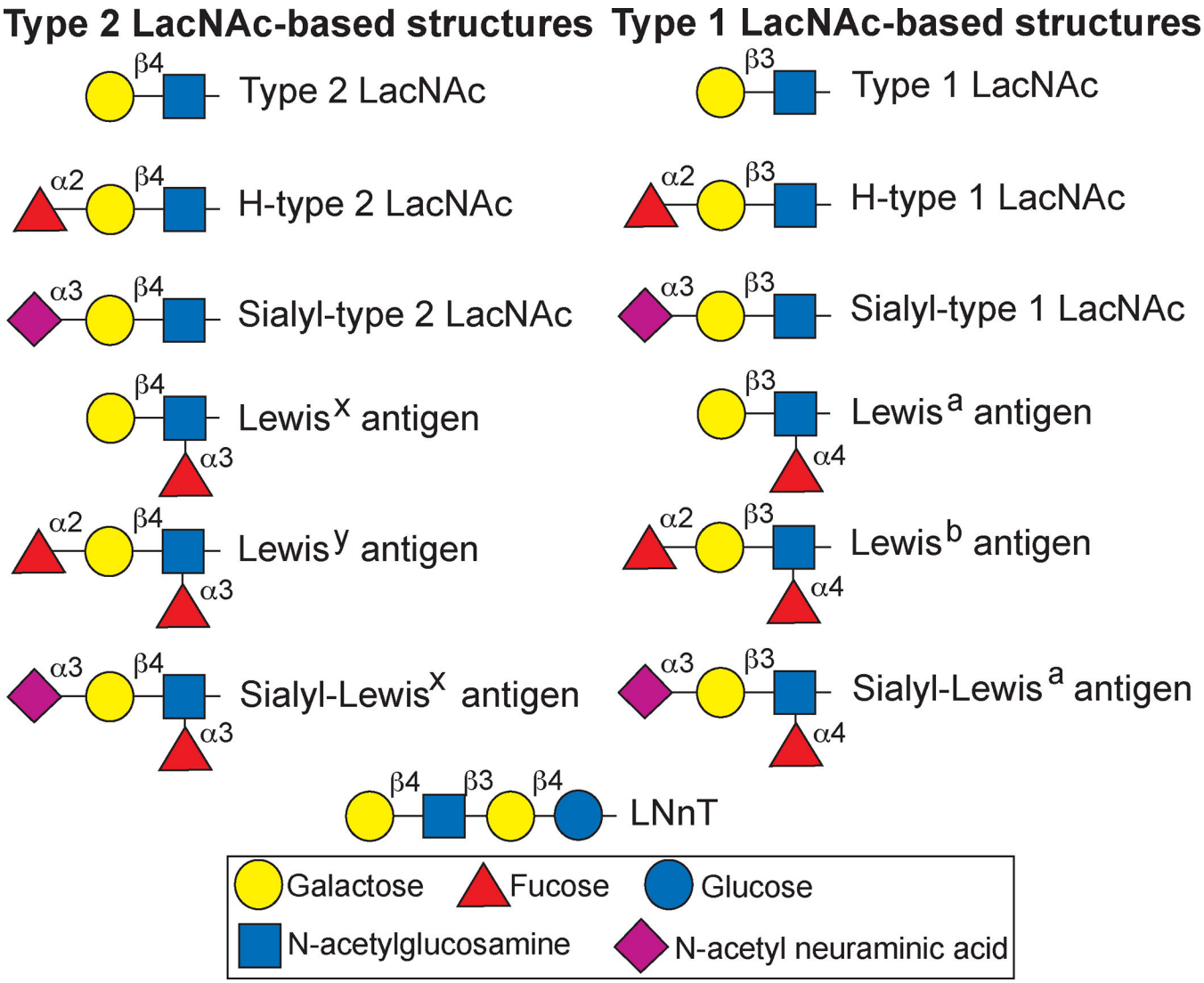


Fig. 1 |. Modified type 1 and type 2 LacNAc structures.

Representative type 2 and type 1 LacNAc terminal structures and additional modifications are shown, including H antigen, Lewis antigen and sialylated Lewis antigen structures (and an extended type 2 lacto-*N*-neotetraose (LNnT) structure). The box contains the legend for monosaccharide cartoon representations.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

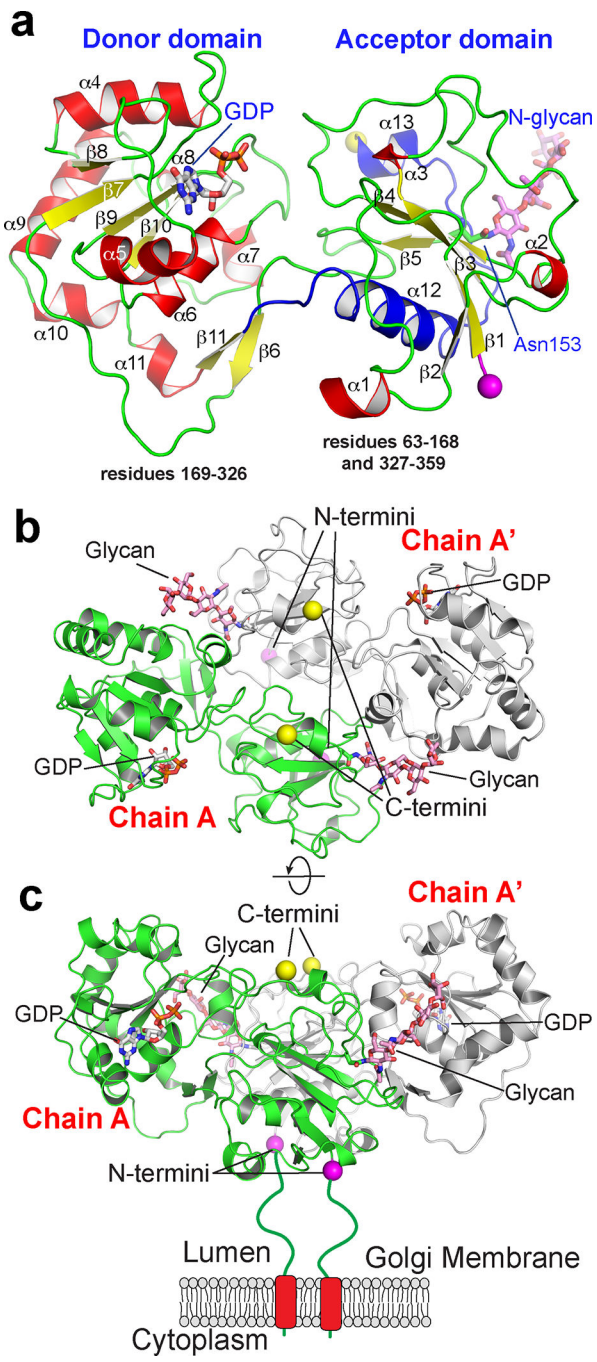


Fig. 2 |. Structure of FUT9 and homodimerization.

a, The GT-B domain structure of FUT9 with secondary structures numbered from the N terminus (magenta sphere). Helices (red), β-sheets (yellow) and loops (green) for the acceptor domain (residues 63–168, 327–359) and donor domain (residues 169–326) along with the bound GDP (white sticks) and N-glycan (pink sticks) attached to Asn153 are shown. The C-terminal (yellow sphere) segment (residues 327–359) extending from the donor domain into the acceptor domain is colored blue. **b**,**c**, The homodimeric structure of FUT9 viewed down the dyad axis (**b**) with chain A colored in green and chain A′ colored

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

in gray. The side view cartoon representation of the biological dimer (**c**) is shown after 90° rotation as a transmembrane, Golgi-localized enzyme in vivo with a 30-amino acid linker 'stem region' (green line) between the transmembrane span (residues 11–32, red ovals) and the catalytic domain (residues 63–359). GDP, glycan and the N- and C-terminal residues are displayed as in **a**.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

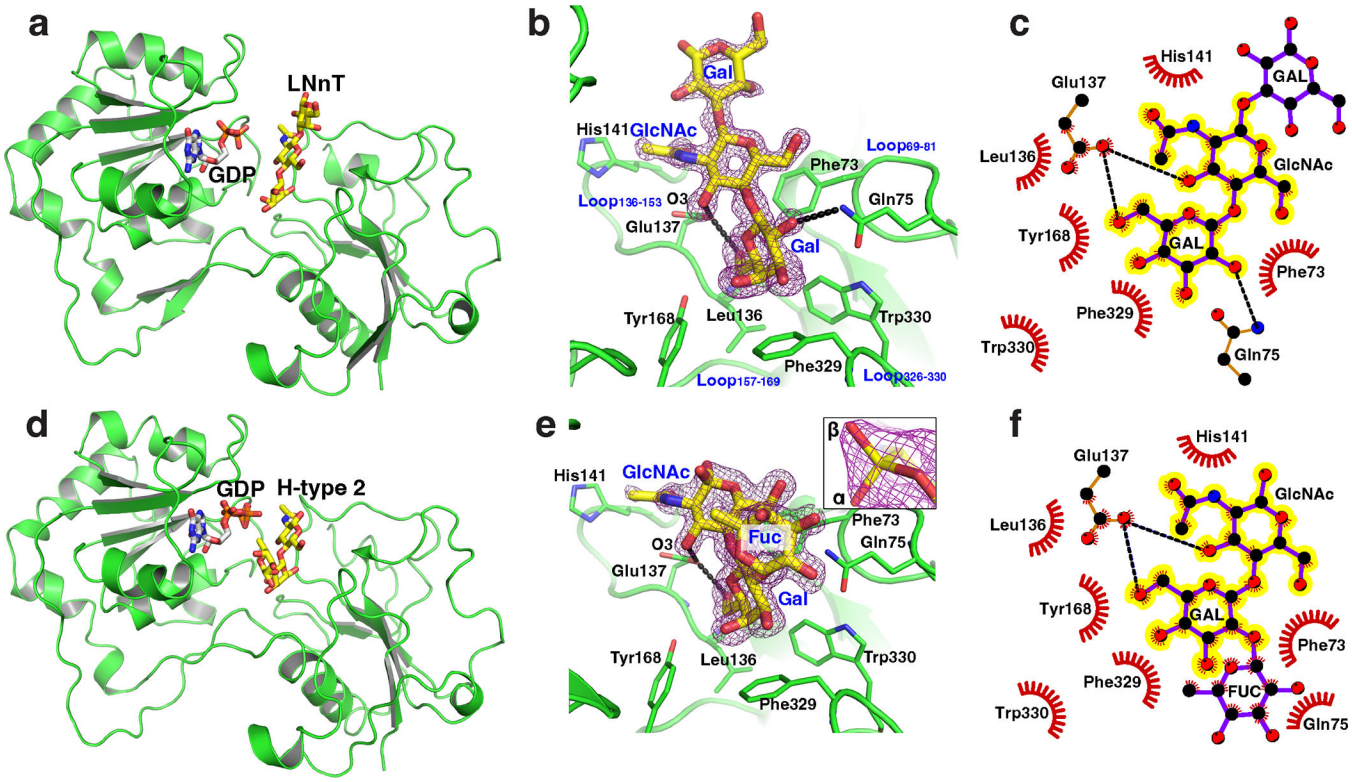


Fig. 3 |. FUT9 acceptor interactions.

a, Cartoon representation of FUT9 with bound GDP (white sticks) and LNnT (yellow sticks) showing the binding of GDP to the donor-binding domain and LNnT bound to the cleft between the donor and acceptor domains. **b**, Difference density map (1.37 Å, contoured at 0.53 $e^{-}/Å^{3}$ (3.0σ)) for LNnT (yellow) in the acceptor-binding site. The map was calculated after omitting LNnT from the refined coordinates and subjecting the model to simulated annealing. The reducing end Glc was disordered and not modeled. Hydrogen bonds (black dashes) are shown. **c**, Ligplot of LNnT (LacNAc highlighted yellow) showing packing interactions (red feathers) and hydrogen bonds (black dashes). **d**, Cartoon representation of FUT9 with bound GDP and H-type 2 acceptor in the same orientation and coloring as in **a**. **e**, Difference density map (1.40 Å, contoured at 0.46 $e^{-}/Å^{3}$ (3.0σ)) for H-type 2 in the acceptor-binding site (colored and calculated as in **b**). Both α- and β-anomers of the reducing terminal GlcNAc are shown (inset). Of note is an alteration in hydrogen bonding for Gln75 in the H-type 2 complex. The hydrogen bond to the Gal O2 hydroxyl in the LNnT complex has been replaced by a very weak hydrogen bond (3.4 Å) with the O2 of the fucose, which is modeled as a van der Waals/electrostatic interaction in the Ligplot. **f**, Ligplot of H-type 2 interactions in the acceptor-binding site (colored and LacNAc highlighted as in **c**).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

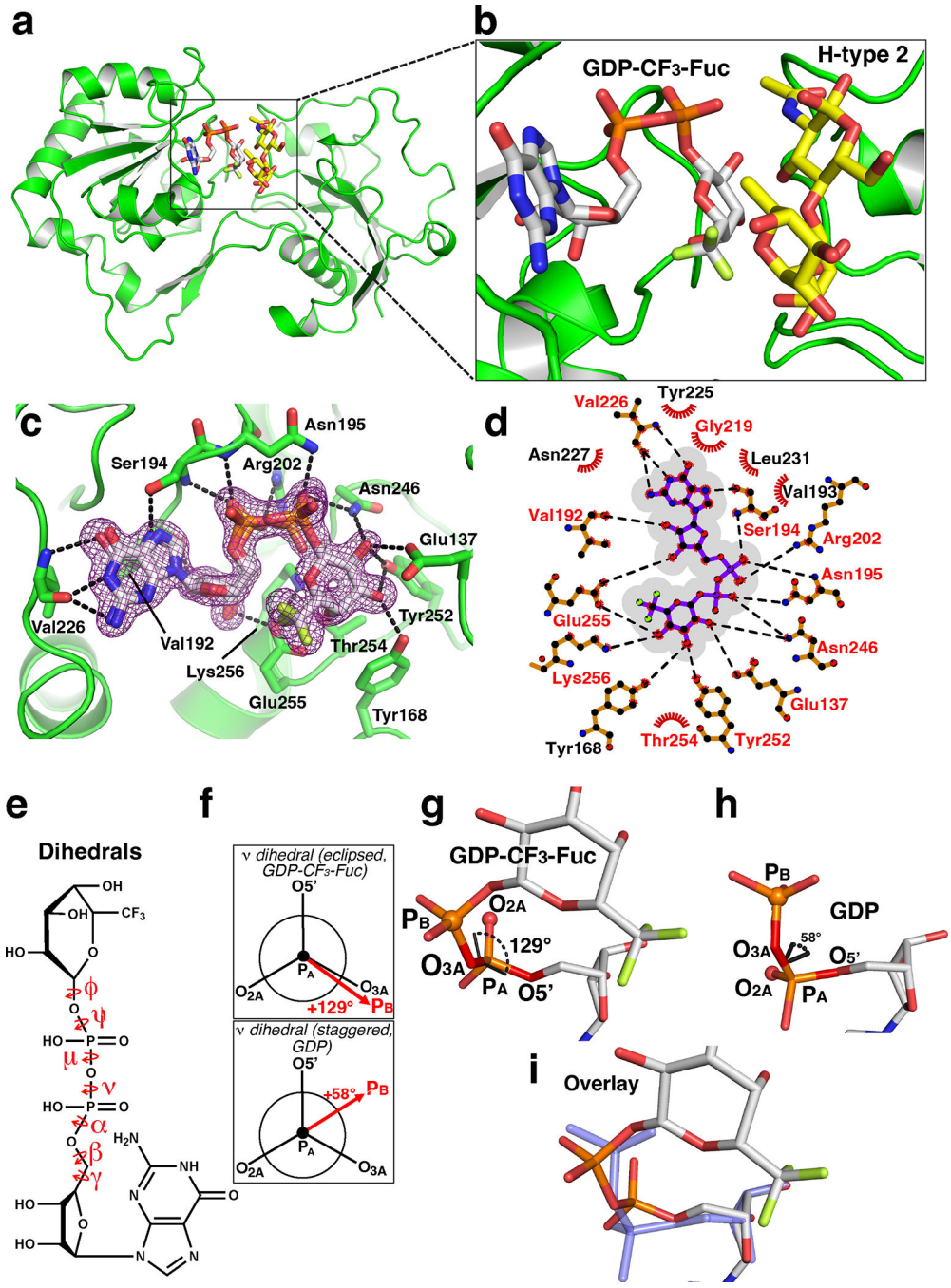


Fig. 4 |. FUT9 donor interactions and hydrolysis.

a, Cartoon representation of FUT9 with bound GDP-CF3-Fuc (white sticks) and H-type 2 (yellow sticks) in the same orientation and coloring as Fig. 3a. **b**, Zoom-in on the structure of the bound FUT9–GDP-CF3-Fuc–H-type 2 complex. **c**, Difference density map for GDP-CF_3-Fuc (1.40 Å, contoured at 0.46 $e^{-}/Å^3$ (3.0σ)) calculated after omitting GDP-CF_3-Fuc from the refined model and subjecting the coordinates to simulated annealing. The electron density reveals a mixture of two diphosphate conformations corresponding to the intact and hydrolyzed donor. Hydrogen bonds are shown as black dashes. **d**, Ligplot of GDP-CF_3-

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Fuc (highlighted gray) showing packing interactions (red feathers) and hydrogen bonds (black dashes). **e**, Dihedral angles of the donor diphosphate in GDP-CF_3-Fuc. **f**, Newman projections of the ν dihedral angle for GDP-CF_3-Fuc (top) and GDP (bottom). **g**,**h**, The ν dihedral angle (cyan bonds, black arc) in GDP-CF_3-Fuc (**g**) and GDP (**h**). **i**, Superposition of GDP-CF_3-Fuc and GDP (slate) to illustrate the conformational change.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

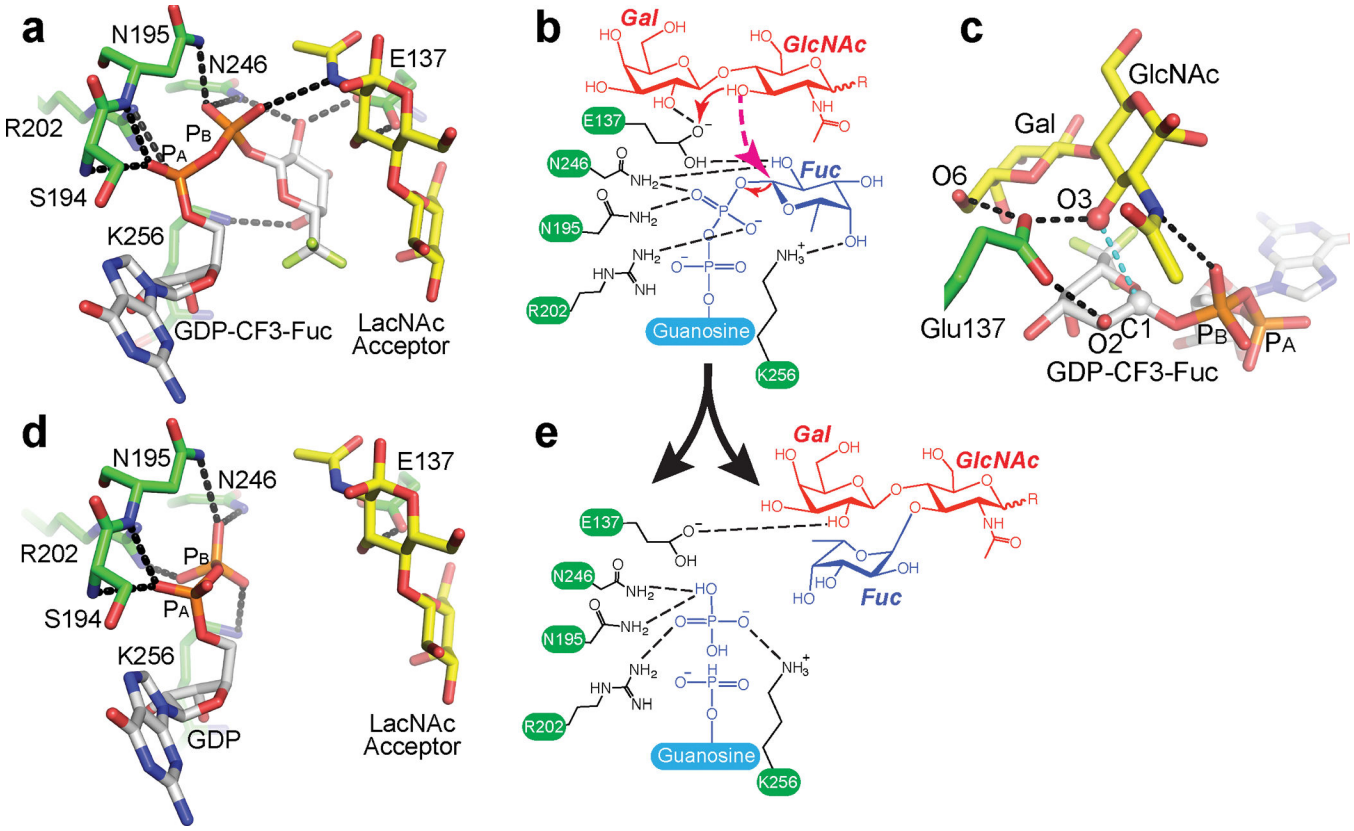


Fig. 5 |. Conformational changes during catalysis.

a, Hydrogen bonding pattern (black dashes) of the diphosphate unit in the intact donor, GDP-CF3-Fuc (pale gray sticks). FUT9 (green sticks) and acceptor H-type 2 (yellow sticks) are shown. **b**, FUT9 reaction mechanism indicating bound GDP-Fuc donor (blue), LacNAc acceptor (red) and FUT9 residues (green). Deprotonation of the acceptor GlcNAc O3 hydroxyl by the E137 catalytic base (red arrow) leads to nucleophilic attack on the Fuc anomeric center (magenta). **c**, The nucleophilic attack of the GlcNAc C3 hydroxyl on the Fuc C1 atom (teal dashes). The hydrogen bonding pattern of Glu137 and deprotonation of the acceptor nucleophile are depicted with black dashes. **d**,**e**, The relaxation of the diphosphate following transfer results in hydrogen bonds with the same residues, albeit with different atoms on the GDP (**d**), and introduces a new salt bridge between P_B and Lys256 (**e**).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

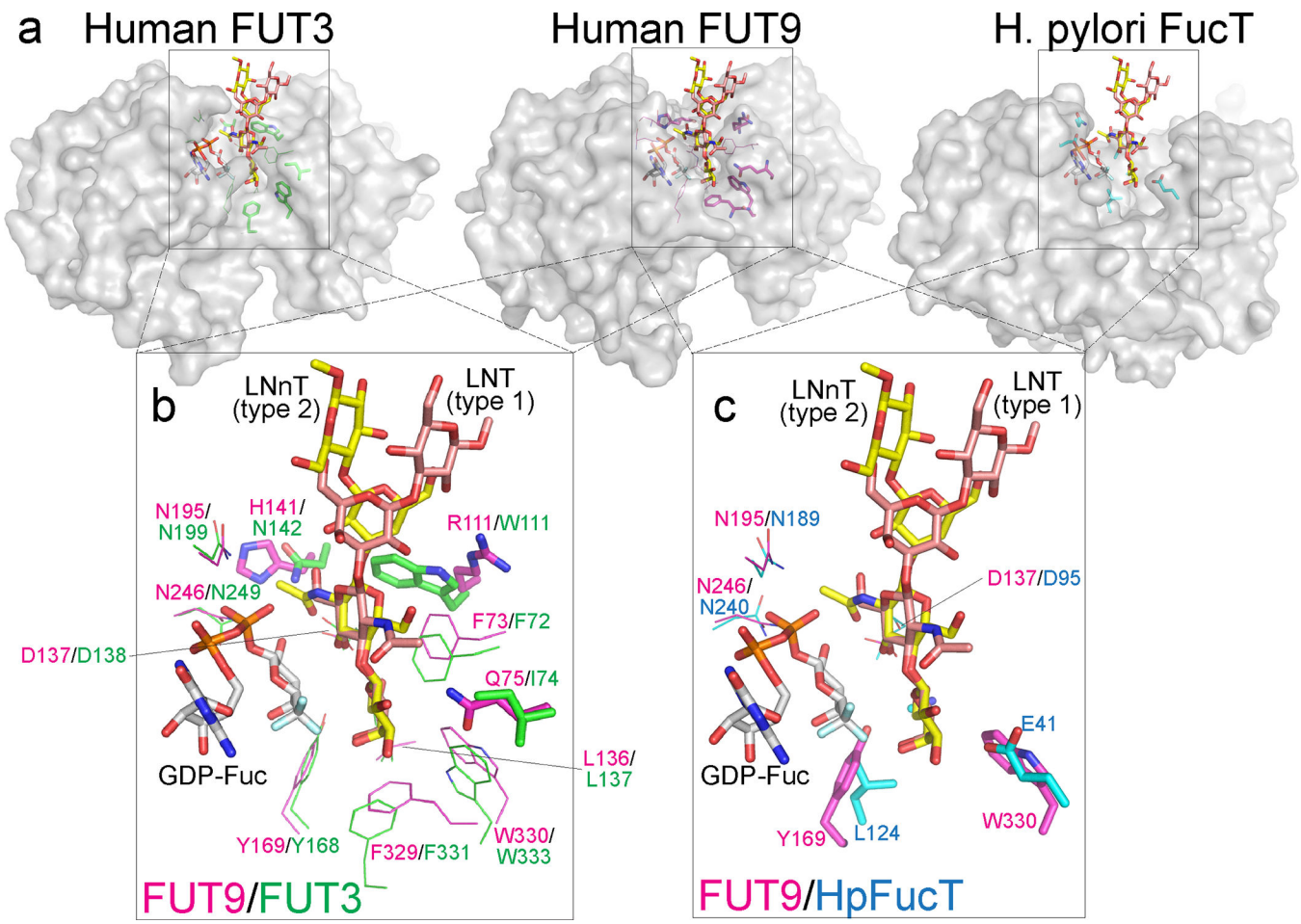


Fig. 6 |. Comparison of modeled type 2 and type 1 chain acceptor complexes for FUT3, FUT9 and HpFucT.

a, The structures of FUT3, FUT9 and HpFucT are shown in gray surface representations with modeled acceptor structures of LNnT (type 2 chain, yellow sticks) or LNT (type 1 chain, salmon sticks) and modeled GDP-CF3-Fuc donor (white sticks). Modeling of bound acceptors was achieved by structural alignment of the AlphaFold[49] modeled structure of FUT3 and HpFucT[29] with the FUT9–GDP-CF3-Fuc–H-type 2 complex. Energy-minimized structures of LNnT and LNT were generated using the Carbohydrate Builder module in Glycam[50]. The acceptor-bound complexes were then generated by superimposing the nonreducing terminal Gal residue from the respective structures to the equivalent Gal residue in the FUT9–GDP-CF3-Fuc–H-type 2 complex. Residues within 4 Å of the respective acceptor are shown either as thin lines for structurally equivalent residues or as stick representation where equivalently positioned residues were different (magenta residues for FUT9, green for FUT3 and cyan for HpFucT). **b**,**c**, Zoomed-in overlays of the FUT9 and FUT3 (**b**) or FUT9 and HpFucT (**c**) are shown as aligned modeled complexes with residues that are structurally equivalent or distinct in each of the respective complexes.