

Structural basis for substrate specificity and catalysis of α 1,6-fucosyltransferase

Ana García-García^{1,9}, Laura Ceballos-Laita ^{1,9}, Sonia Serna^{2,3}, Raik Artschwager^{2,3}, Niels C. Reichardt^{2,3,4}, Francisco Corzana ⁵ & Ramon Hurtado-Guerrero ^{1,6,7,8} 

Core-fucosylation is an essential biological modification by which a fucose is transferred from GDP- β -L-fucose to the innermost N-acetylglucosamine residue of N-linked glycans. A single human enzyme α 1,6-fucosyltransferase (FUT8) is the only enzyme responsible for this modification via the addition of an α -1,6-linked fucose to N-glycans. To date, the details of substrate recognition and catalysis by FUT8 remain unknown. Here, we report the crystal structure of FUT8 complexed with GDP and a biantennary complex N-glycan (**GO**), which provides insight into both substrate recognition and catalysis. FUT8 follows an S_N2 mechanism and deploys a series of loops and an α -helix which all contribute in forming the binding site. An exosite, formed by one of these loops and an SH3 domain, is responsible for the recognition of branched sugars, making contacts specifically to the α 1,3 arm GlcNAc, a feature required for catalysis. This information serves as a framework for inhibitor design, and helps to assess its potential as a therapeutic target.

¹Institute of Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Mariano Esquillor s/n, Campus Rio Ebro, Edificio I+D, Zaragoza, Spain. ²CIC biomaGUNE, Paseo Miramón 182, San Sebastián, Spain. ³Basque Research and Technology Alliance (BRTA), Paseo Miramón 182, San Sebastian, Spain. ⁴CIBER-BBN, Paseo Miramón 182, San Sebastian, Spain. ⁵Departamento de Química, Universidad de La Rioja, Centro de Investigación en Síntesis Química, E-26006 Logroño, Spain. ⁶Copenhagen Center for Glycomics, Department of Cellular and Molecular Medicine, School of Dentistry, University of Copenhagen, Copenhagen, Denmark. ⁷Laboratorio de Microscopías Avanzada (LMA), University of Zaragoza, Mariano Esquillor s/n, Campus Rio Ebro, Edificio I+D, Zaragoza, Spain. ⁸Fundación ARAID, Zaragoza, Spain. ⁹These authors contributed equally: Ana García-García, Laura Ceballos-Laita. email: rhurtado@bifi.es

Core-fucosylation is an essential biological modification of the *N*-glycan core in eukaryotes (except for plants and fungi) that is performed by the inverting α 1,6-fucosyltransferase (FUT8)¹. It is the most common type of fucose modification and occurs in the Golgi apparatus. FUT8 transfers a *n* 1-fucose residue from GDP- β -L-fucose (GDP-Fuc) onto the innermost GlcNAc of *N*-glycan to form an α 1,6-linkage². Gene knockout of *FUT8* in mice led to early postnatal death, severe growth retardation and emphysema-like changes in the lung, and revealed that this modification is crucial for the activation of growth factor receptors^{1,3}. FUT8 is also involved in cell adhesion and cell migration processes by influencing turnover and expression levels of E-cadherin⁴ and the activity of α 3/ β 1 integrin⁵. Core-fucose also effects *N*-glycan conformation by favoring an extended orientation of the α 1,6 arm⁶. This can lead to subtle changes in glycan recognition by lectin receptors⁷. However, core-fucosylation is likely most known for its capacity to adjust the humoral immune response by altering antibody-dependent cellular cytotoxicity (ADCC). Thus, different strategies for reducing core-fucosylation in therapeutic monoclonal antibodies have been developed to enhance ADCC⁸. Notably, FUT8 is up-regulated in numerous types of cancer, suggesting that blocking its activity could be a promising strategy for improving anti-tumor immune responses⁹.

The acceptor specificity of FUT8 requires the presence of a terminal GlcNAc moiety on the α 1,3 arm of the *N*-glycan but shows a higher degree of flexibility on the α 1,6 arm¹⁰. This substrate preference does not necessarily demand the presence of a peptide/protein¹⁰. On the contrary, FUT8 also fucosylates high mannose *N*-glycans lacking a terminal GlcNAc moiety on the α 1,3 arm. However, in these cases a peptide/protein moiety attached to the first GlcNAc via an *N*-glycosidic linkage¹¹ is necessary.

The crystal structure of the human FUT8 (*HsFUT8*) apo form has been previously reported revealing that FUT8 is composed of a multi-domain enzyme that contains an *N*-terminal coiled-coil domain, a catalytic domain, which adopts a GT-B fold, and a C-terminal SH3 domain¹². However, the lack of solved ternary complexes has impeded to obtain mechanistic insights into the glycosyl transfer reaction or reveal the molecular basis of the requirement for a terminal GlcNAc moiety on the α 1,3 arm for optimal catalysis. Here, we have captured a ternary complex formed between the human FUT8 (*HsFUT8*), GDP and the biantennary complex *N*-glycan **G0** (see below) by X-ray crystallography, uncovering the molecular basis of FUT8 catalysis and recognition of acceptor substrates.

Results

Core-fucosylation of synthetic *N*-glycans by *HsFUT8*. *HsFUT8* was successfully cloned, expressed and purified to homogeneity (Methods) to enable functional and structural characterization. The activity of *HsFUT8* on 5-aminopentyl **G0** was monitored by mass spectrometry that showed a mass increase of 146 Da corresponding to the incorporation of a fucose moiety into the glycan in the presence of GDP-Fuc (Supplementary Fig. 1). Then, we evaluated the *HsFUT8* acceptor substrate specificity on a microarray of synthetic *N*-glycans by expanding an earlier array version used by us in the study of *C. elegans* FUT8 (*CeFUT8*)^{13,14} (Fig. 1, Supplementary Table 1). The *HsFUT8* showed identical substrate acceptor preferences to the reported ones for *CeFUT8* (Fig. 2), and verified prior reports describing the importance of a terminal GlcNAc moiety in the α 1,3 arm as a requirement for catalysis and a great promiscuity towards modifications on the α 1,6 arm¹⁰.

Architecture of the *HsFUT8*-GDP-G0** complex.** Further, we successfully obtained C2 crystals of *HsFUT8* in a complex with

GDP and **G0** glycan (Table 1). The resulting crystals allowed us to solve the structure at high resolution (1.95 Å) and interpret the density map (Table 1). The asymmetric unit (AU) of C2 crystals contains two molecules of *HsFUT8* that are arranged as dimers with two molecules from the neighboring AU by direct contact between their corresponding *N*-terminal coiled-coil domains (Fig. 3). The root mean square deviation (RMSD) between both molecules belonging to chain A and B in the AU is 0.44 Å on 452 equivalent Ca atoms (Fig. 4; hereafter we will only discuss molecule A because it does not contain any disordered regions). As expected, the analysis of the *HsFUT8* structure with the DALI server¹⁵ revealed structural homology to two fucosyltransferases, namely the *Arabidopsis thaliana* FUT1 (*AtFUT1*; PDB entries 5KX6, 5KOP, 5KOR, 5KWK and 5KOE^{16,17}) and the human and *C. elegans* PoFUT2 (*HsPoFUT2* and *CePoFUT2*; PDB entries 4AP5, 4AP6 and 5FOE^{18–20}). Although FUT8 is very distant to *AtFUT1* and *HsPoFUT2*/*CePoFUT2* in terms of acceptor substrates, the server rendered good scores implying that they superimposed fairly well (RMSDs of ~3.2 and ~5.0 Å between *AtFUT1* and *HsFUT8*, and *AtFUT1* and *HsPoFUT2*/*CePoFUT2* crystal structures, respectively; the superimposed residues ranged from 429 to 459 residues).

The structure of *HsFUT8* complexed to GDP and **G0** depicts the three aforementioned domains and how they are connected with each other. While the catalytic domain is connected to the coiled-coil domain by α 3, the SH3 domain is connected to the catalytic domain by the loop β 10– β 11 (Fig. 4a, Supplementary Fig. 2). While in general, densities for GDP and the rest of the **G0** molecule are very well defined, we only observe partial density for the GlcNAc moiety on the α 1,6 arm (**G^{G0}**). Of special note is the *anti- ψ* conformation for the core-chitobiose GlcNAc moieties (glycosidic linkage involving units **A^{G0}–B^{G0}**) of the *N*-glycan in the bound state (see below). This conformer is less stable relative to the typical *syn- ψ* conformation found in solution²¹ (see inset in Fig. 4a). In addition, the Man α 1-6Man linkage of **G0** adopts an ‘extended gg’ conformer²², with ψ and ω values of 174.7° and 50.4°, respectively. The structure also reveals that GDP is partly buried and located exclusively in the catalytic domain, while the inner part and the α 1,3/ α 1,6 arms of **G0** are located in the catalytic domain and the exosite formed by the loop β 10– β 11 and SH3 domain, respectively (Fig. 4b). The presence of exosites for intimate binding to *N*-glycans was also observed in distant glycosyltransferases (GTs) and glycosyl hydrolases such as MGAT2 and MAN2A1, respectively²³. Our structure also exhibits two striking conformational changes relative to the apo form (RMSD of 0.56 Å on 436 equivalent C α s; Fig. 4c). Residues 428–444 (formed by a loop and α 10) experience a large conformational change (the largest distance, close to 22 Å, was observed for Gly437 in both forms) that allows the partial closure of the GDP-Fuc-binding site and which protects the more hydrophobic part of the molecule (the guanosine moiety) from the solvent. Significantly, during the transition from the apo to the substrate-bound forms, a rearrangement of the α 10 takes place in which α 10^{apo-form} (residue 436–442) becomes unstructured and another α 10, comprising residues 430–435, is formed. Furthermore, the loop β 6– α 8 (residues 365–378) is partly disordered in the apo form and suffers a conformational change in the presence of both ligands, leading to an ordered loop that contributes to the formation of both the GDP-Fuc and the acceptor binding sites (Fig. 4c).

The sugar nucleotide-binding site. The *HsFUT8*-binding site is large and set up by the GDP-Fuc and the *N*-glycan binding sites (Fig. 5a, b). The guanine moiety of GDP establishes CH– π interactions with Ala436, Val450 and Val471 while the guanosine

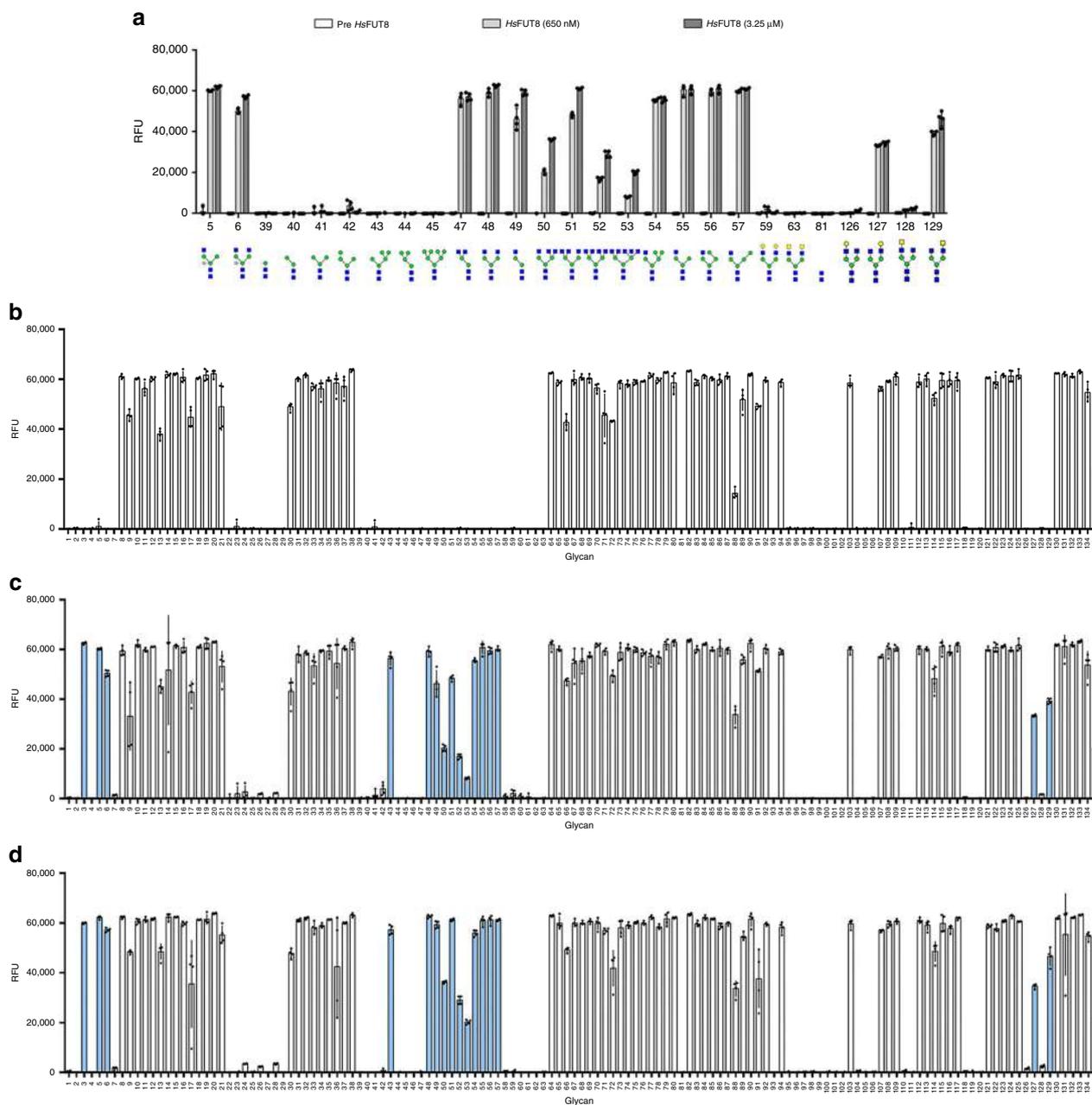


Fig. 2 Activity assay of HsFUT8 on glycan microarray. a *Aleuria aurantia* lectin (AAL-555) binding before and after incubation with recombinant HsFUT8 at two different concentrations (650 nM and 3.5 μM) on a selection of glycan structures. **b** AAL-555-binding profile on a full glycan array before treatment with recombinant HsFUT8. **c** AAL-555-binding assay after treatment with recombinant HsFUT8 at 650 nM enzyme concentration and **d** AAL-555-binding assay after treatment with recombinant HsFUT8 at 3.25 μM enzyme concentration. Histogram bars show the average fluorescence RFU (relative fluorescence units) values for four replicate spots on the array. Error bars represent the standard deviation of the average RFU values on the same microarray. N-glycans that show binding towards AAL after incubation with HsFUT8 are highlighted in blue.

(Fig. 5a), in line with previous SPR data in which the binding of GDP-Fuc was ~40-fold stronger than the binding of G⁰ to the enzyme²⁶ (Fig. 5a). A^{G0} is recognized through hydrogen bonds formed between A^{G0} OH1 and Gly217 backbone, A^{G0} acetamide NH and Asp295 side chain, and A^{G0} OH3 and Asp295/Lys216 side chains. B^{G0} is tethered via interactions formed between B^{G0} OH6 and Glu373 side chain, and B^{G0} acetamide carbonyl group and Gln470 side chain. The unusual *anti-ψ* conformation of A^{G0} and B^{G0}, which is imposed by the enzyme, is a strict requirement for the A^{G0} OH6 fucosylation site facing GDP-Fuc (Fig. 5a). Catalysis can only take place in the *anti-ψ* conformation even though this conformation is less energetically favorable than the solution-phase

syn-ψ conformation (−8.9 versus −9.5 kcal/mol, estimated by MM calculations for the disaccharide containing A^{G0}–B^{G0} units).

The sugars forming the branching part of G⁰ (C^{G0}, D^{G0}, E^{G0}, F^{G0} and G^{G0}) are located in the exosite formed by the loop β10–β11 and the SH3 domain (Fig. 5a). C^{G0} establishes a CH–π interaction with Tyr498 and provides one hydrogen bond between C^{G0} OH4 and Asp495 side chain. D^{G0} is poorly recognized by a lonely hydrogen bond between D^{G0} OH6 and the Glu373 backbone. However, the terminal GlcNAc moiety in the α1,3 arm, E^{G0}, is the most intimately recognized monosaccharide residue of all branched sugars. E^{G0} established three hydrogen bonds that are formed between E^{G0} OH6 and His535 side chain,

Table 1 Data collection and refinement statistics.

HsFUT8 in complex with GDP and G0	
Data collection	
Space group	C2
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	208.04, 68.59, 173.92
α , β , γ (°)	90, 149.90, 90
Resolution (Å)	20–1.95 (2.06–1.95) ^a
<i>R</i> _{merge}	0.133 (1.368)
<i>I</i> / σ <i>I</i>	6.4 (1.4)
Completeness (%)	99.9 (100)
Redundancy	4.4 (4.5)
Refinement	
Resolution (Å)	1.95
No. of reflections	394,687
<i>R</i> _{work} / <i>R</i> _{free}	0.185/0.218
No. of atoms	
Protein	7511
GDP	56
G0	180
Waters	506
Glycerol	102
<i>B</i> -factors (Å ²)	
Protein	38.11
GDP	47.64
G0	44.70
Waters	49.00
Glycerol	63.01
R.m.s. deviations	
Bond lengths (Å)	0.0159
Bond angles (°)	2.0646

One crystal was used to determine the crystal structure. ^aValues in parentheses are for highest-resolution shell.

E^{G0} OH2 and Asp494 backbone, and E^{G0} acetamide NH and Asp495 side chain. In addition, the E^{G0} acetamide methyl group is engaged in a hydrophobic interaction with Ala374 (Fig. 5a). The contacts that E^{G0} establishes with the enzyme could explain the strict requirement for a GlcNAc moiety in this position. This unit improves the binding of the α 1,3 arm, and in turn, provides stabilization of the entire molecule in the bound state. On the contrary, the branching sugars on the α 1,6 arm are far less well recognized, which provides a plausible explanation for the larger promiscuity towards the α 1,6-mannose branch¹⁰. The F^{G0} endocyclic oxygen and F^{G0} OH4 establish two hydrogen bonds with Gln502 side chain and Val531 backbone, respectively, while other potential interactions between G^{G0} and HsFUT8 (G^{G0} carbonyl group and G^{G0} OH3 with Gly501 and Gln502 backbones, respectively) cannot be confirmed due to the less well-defined density in this sugar residue (Figs. 4a and 5a). The addition of a bisecting GlcNAc in G0 would preclude fucosylation by FUT8 because this residue is likely to impose a significant steric hindrance to the complex (Fig. 6), explaining why *N*-glycans with bisecting GlcNAc are not substrates of FUT8 (ref. 27).

Discussion

FUT8 is the only enzyme responsible for core-fucosylation on mammalian *N*-glycoproteins. Despite many years of research on the enzymology and chemistry of FUT8, its reaction mechanism remained elusive and the requirement for a terminal GlcNAc moiety on the α 1,3 arm of the *N*-glycan for optimal fucosylation was not well understood. The structure of the HsFUT8 apo form provided the architecture of this enzyme, which is formed by three different domains¹². However, this structure did not offer any clues for the reaction mechanism or substrate recognition. Here, we have determined the crystal structure of a functional HsFUT8 in complex with GDP and G0. Intriguingly, the structure

reveals that residues 436–442, in which one loop and α 10 are present, and the loop β 6– α 8 suffer large conformational changes in the presence of the ligands forming the GDP-Fuc and acceptor binding sites. Though conformational changes of loops have been visualized in inverting and retaining GTs^{24,28,29}, the motion here of the loop β 6– α 8 in the presence of ligands is unique because this loop brings in key residues that recognize both GDP and G0. In particular, Arg365, Lys369 and Glu373 contribute not only to the binding but also to catalysis, with Glu373 acting as the catalytic base for A^{G0} OH6. This is compatible with an S_N2 single-displacement reaction mechanism. Generally, GT-A or GT-B inverting GTs following an S_N2 inverting mechanism contain an amino acid acting as a catalytic base²⁴ (typically Glu, Asp, or His). However, in most structures, these residues acting as catalytic residues are already located in the binding site in the absence of ligands²⁴. Hence and to our knowledge, FUT8 exemplifies a unique case in which the catalytic base is brought to the binding site in the presence of ligands (note that the distance of Glu373 C α s of the apo and substrate-bound form is ca. 18 Å; Fig. 7).

Previously and based on computational studies, HsFUT8 was suggested to follow an S_N2 mechanism in which the β -phosphate was acting as the catalytic base²⁶. Similarly, the β -phosphate acting as the catalytic base was proposed for the inverting CePoFUT1²⁵, though in this case an S_N1 mechanism was suggested. In both enzymes, there was no basic residue present in the active site that could act as an assisting base. However, and based on our crystal structure, it is now clear that the authors could not predict the motion of the loop β 6– α 8, which brings the essential catalytic base Glu373. This demonstrates that it is imperative to obtain ternary complexes for these enzymes to elucidate their catalytic mechanism. Our S_N2 inverting mechanism is also reinforced by the previous elucidation of ternary complexes for the closest structural homologs, AtFUT1 (ref. 17) and CePoFUT2 (ref. 18). For both enzymes, residues such as Asp and Glu, respectively, were also acting as the catalytic bases.

At the level of the *N*-glycan-binding site, four major structural features are responsible of the selectivity and recognition of HsFUT8 on *N*-glycans: (a) the less stable *anti-ψ* conformation for the core-chitobiose GlcNAc moieties of the *N*-glycan in the bound state is a pre-requisite for core-fucosylation, since it provides the A^{G0} OH6 in close contact to GDP-Fuc; (b) the CH– π interaction between Tyr498 and C^{G0} is likely crucial for a right orientation of the branching sugars; (c) the existence of an exosite formed by the loop β 10– β 11 and SH3 domain is responsible for branching sugar recognition and, in particular, for the α 1,3 arm GlcNAc binding. SH3 domains are small protein domains of about 60 amino acid residues present in a large number of proteins and well-known for driving protein–protein interactions³⁰. The fact that residues from the HsFUT8 SH3 domain recognize sugar moieties implies the broad versatility of these domains in recognition of different molecules; and (d) the most intimately recognition for the α 1,3 arm GlcNAc explains why FUT8 requires the presence of a terminal GlcNAc moiety on the α 1,3 arm of the *N*-glycan for optimal fucosylation. This ligand-binding model also explains our own data using HsFUT8 on a microarray of synthetic *N*-glycans and prior reports^{10,13,14}. However, our structure does not offer clues on how FUT8 fucosylates less optimally high mannose *N*-glycans linked to glycoproteins and peptides that lack a terminal GlcNAc on the α 1,3-branch and which will require further structural studies.

In conclusion, the structure reported here shows how FUT8 recognizes *N*-glycans and provides information on the enzymatic mechanism. Our structure may assist in the development of inhibitors with a potential value for cancer therapy or as an alternative for the production of non-fucosylated antibodies with enhanced ADCC.

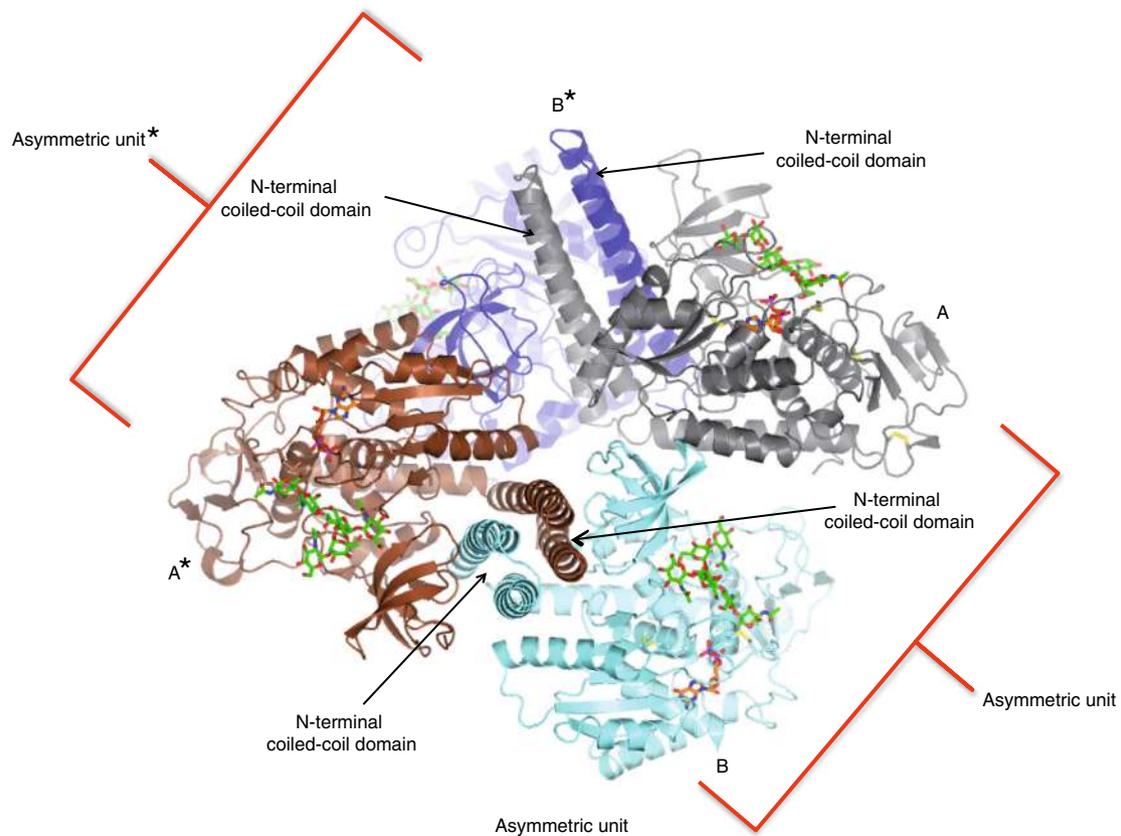


Fig. 3 Dimeric structure of *HsFUT8*. One *HsFUT8* molecule (A in gray) from one asymmetric unit forms a dimeric structure with another molecule (B* in blue) from the other asymmetric unit* through their corresponding coiled-coil domains.

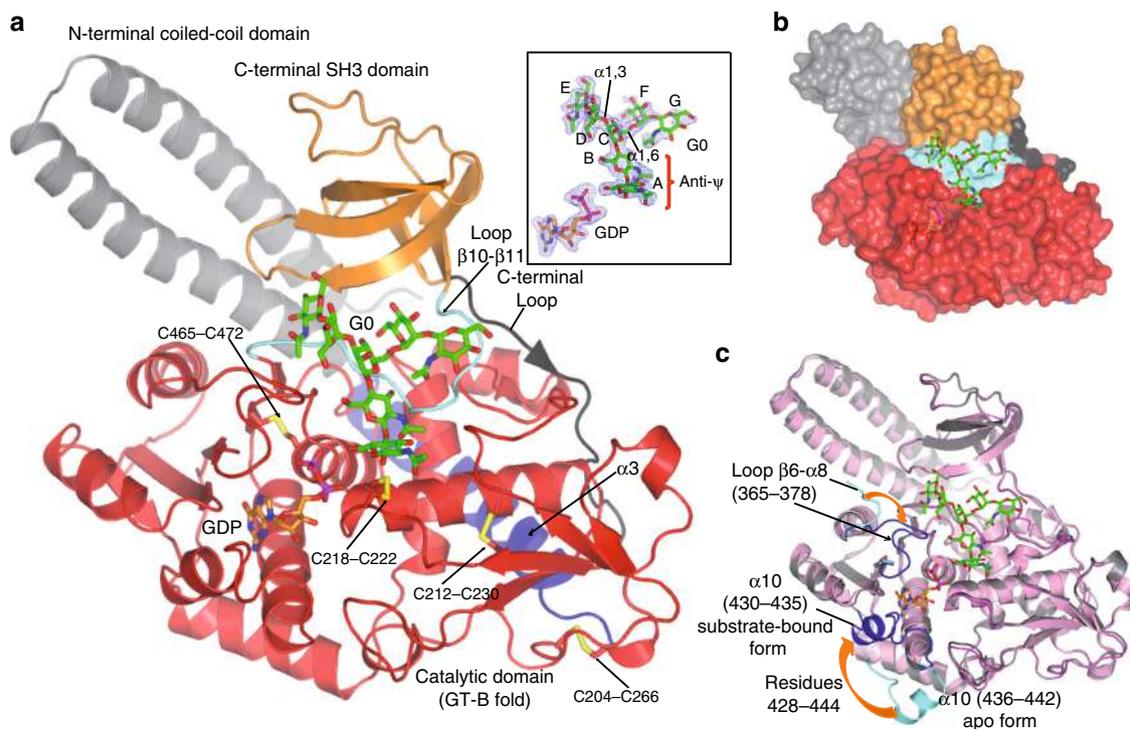


Fig. 4 Overall structure of *HsFUT8* complexed to GDP and GO. **a** Ribbon structure of *HsFUT8* with GDP (orange carbon atoms) and GO (green carbon atoms). The coiled-coil, catalytic and SH3 domains are colored in gray, red and orange, respectively. The interdomain $\alpha 3$ and loop $\beta 10$ - $\beta 11$ are colored in blue and aquamarine, respectively. Disulfide bridges are indicated as yellow sulfur atoms. The C-terminal loop is colored in black. (inset) Electron density maps are F_O - F_C (blue) contoured at 2.2σ for GDP and GO. The labels of the sugar units (A-G) are also shown. **b** Surface representation of the *HsFUT8*-GDP-GO complex. **c** Superposition of *HsFUT8* apo (pink) and substrate-bound (gray) forms. This figure also depicts the conformational changes visualized in residues 365-378 and 428-444. These regions are represented as aquamarine (apo form) and blue (substrate-bound form).

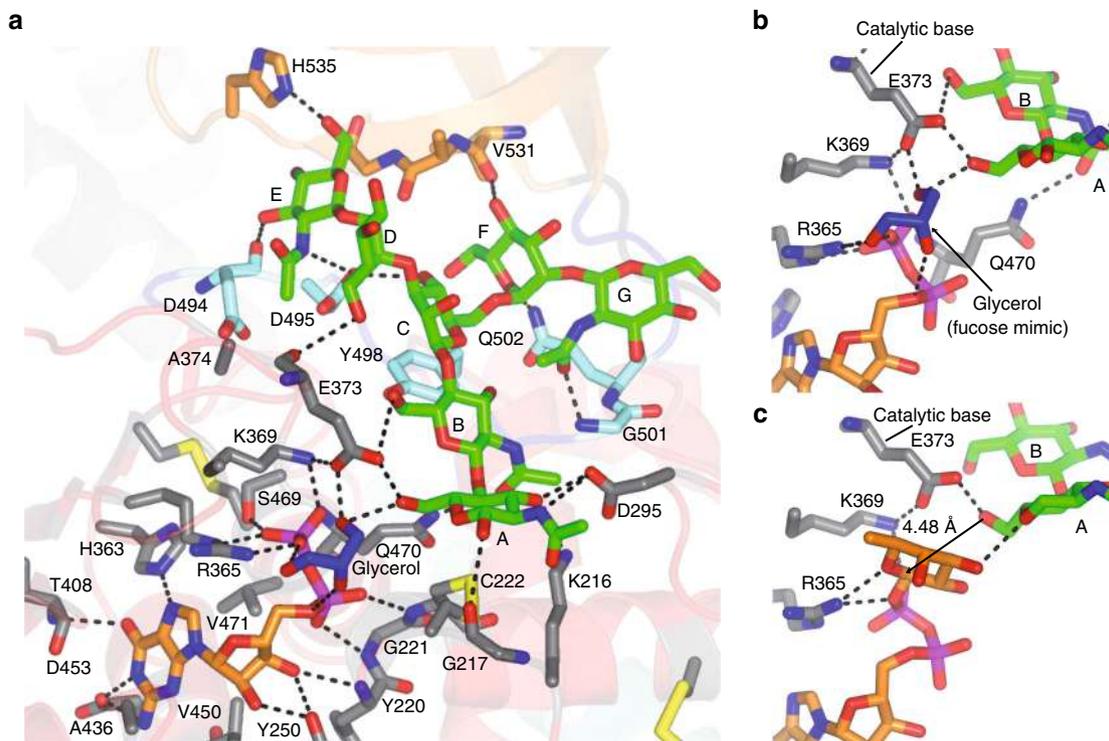


Fig. 5 Structural features of GDP-Fuc and GO binding sites. **a** Complete GDP-Fuc and GO binding sites of the HsFUT8-GDP-GO complex. The residues forming the GDP-Fuc/A^{GO}/B^{GO} and C^{GO}/D^{GO}/E^{GO}/F^{GO}/G^{GO} binding sites are depicted as gray and aquamarine/orange carbon atoms, respectively. GDP, GO and glycerol are shown as orange, green and blue carbon atoms, respectively. Hydrogen bond interactions are shown as dotted black lines. **b, c** Close-up view of the binding site region of the HsFUT8-GDP-GO and HsFUT8-GDP-Fuc-GO complexes showing the essential residues (Arg365, Lys369 and Glu373) that are major players in the plausible S_N2 single-displacement reaction mechanism. Note the proximity and the orientation of the A^{GO} OH6 to the anomeric carbon (4.48 Å) which is compatible with the inversion of the configuration during the reaction.

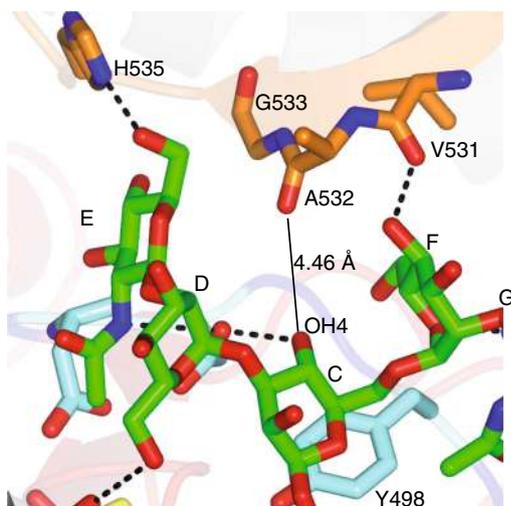


Fig. 6 The bisecting GlcNAc in N-glycans. N-glycans carrying a bisecting GlcNAc are not substrates of HsFUT8 likely due to a steric clash between the bisecting GlcNAc with Ala532 (distance of 4.46 Å between the Ala532 backbone and C^{GO} OH4). Note that the bisecting GlcNAc would be bound to C^{GO} via the C^{GO} OH4, and it is this GlcNAc who would have a steric clash with Ala532. It is also likely that the addition of the bisecting GlcNAc causes conformational changes to the neighboring branching sugar moieties, which could contribute to disrupt the binding to HsFUT8.

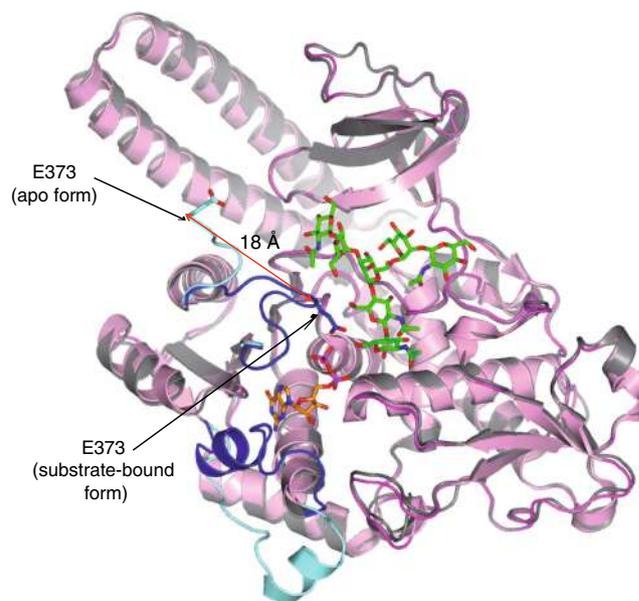


Fig. 7 Superposition of HsFUT8 apo (pink) and substrate-bound (gray) forms. This figure also depicts the conformational changes visualized in Glu373 (in aquamarine and blue for the apo and substrate-bound form, respectively).

Methods

Glycan preparation. For the preparation of the biantennary complex A2 N-glycan (named here as **G0**), sialyl glycopeptide (SGP) predominantly glycosylated with **A2G2S2** glycan was isolated from hen's egg yolk according to a published procedure³¹ and further enzymatically and chemically processed. Treatment of the glycopeptide with acetic acid at 80 °C removed sialic acid residues³² and deglycosylation with PNGase F (Asparica Glycomics) produced a mix of mono and bis-galactosylated glycans³³. The obtained mixture (110 mg, 66.8 μmol) was dissolved in 50 mM citric phosphate buffer (8 mL, pH = 5). *Aspergillus niger* β-galactosidase (8.9 U/mg, 14.2 mg, in 1260 μL buffer) and an aqueous solution of 0.02% Na₂S₂O₃ were added. The reaction mixture was incubated at 37 °C and the reaction progress monitored via MALDI-TOF MS. After complete degalactosylation, MeOH (5 mL) was added and the reaction mixture centrifuged at 4500 r.p.m. MeOH was removed under reduced pressure and the supernatant was loaded onto a Bond elute graphite cartridge (10 g) and eluted with H₂O (250 mL), 25% MeOH/H₂O (500 mL), 50% MeOH/H₂O (250 mL). The collected fractions of 50% MeOH were analyzed by MALDI-TOF MS and fractions containing pure product were pooled and lyophilized. The pure product **G0** was obtained as a white solid (69.1 mg, 52.5 μmol, 79%). **¹H NMR** (500 MHz, D₂O, 298 K): δ 5.10 (s, 1H, αH-1_{GlcNAc-1}), 5.02 (s, 1H, H-1_{αMan-2}), 4.83 (s, 1H, H-1_{αMan-3}), 4.68 (s, 1H, H-1_{βMan-1}), 4.62–4.59 (m, 1H, βH-1_{GlcNAc-1}), 4.54–4.50 (m, 1H, H-1_{βGlcNAc-2}), 4.46 (d, *J* = 8.4 Hz, 2H, H-1_{βGlcNAc-3}), H-1_{βGlcNAc-4}), 4.16 (s, 1H, H-2_{αMan-2}), 4.10 (s, 1H, H-2_{αMan-3}), 4.02 (s, 1H, H-2_{αMan-1}), 3.90–3.31 (m, 39H), 1.99 (s, 3H, CH₃_{GlcNAc}), 1.96 (s, 6H, CH₃_{GlcNAc}), 1.95 (s, 3H, CH₃_{GlcNAc}). **¹³C NMR** (126 MHz, D₂O, 298 K): δ 174.8, 174.7, 174.6 (4 × CO), 101.3 (C1_{βGlcNAc-2}), 100.4 (C1_{βMan-1}), 99.5 (C1_{βMan-2}, C1_{βGlcNAc-2}, C1_{βGlcNAc-3}), 97.0 (C1_{βMan-3}), 94.8 (βC1_{GlcNAc-1}), 90.6 (αC1_{αGlcNAc-1}), 79.6, 79.5, 79.2, 76.4 (C2_{Man-3}), 76.3 (C2_{Man-1}), 75.8, 74.3, 73.5, 73.4, 73.2, 72.8, 72.4, 72.0, 69.9 (C2_{Man-2}), 69.2, 67.3, 65.7, 61.6, 60.6, 59.9, 55.3 (C2_{GlcNAc-4}), 54.9 (C2_{GlcNAc-3}), 53.6 (C2_{GlcNAc-1}, C2_{GlcNAc-2}), 22.3, 22.2, 21.8 (4 × CH₃). **HRMS** (MALDI-TOF): calcd for C₅₀H₈₄N₄O₃₆Na [M + Na]⁺ 1339.475, found 1339.515.

Activity assay on 5-aminopentyl **G0.** **G0** containing a C5 linker at the anomeric position of the reducing end (5-aminopentyl **G0**³⁴) (1 nmol), guanosine 5'-diphospho-β-L-fucose (GDP-Fuc) (2 nmol) and **HsFUT8** (18 μM) in Tris 25 mM, 150 mM NaCl pH = 7.5 were incubated at room temperature (rt) for 5 h. MALDI-TOF MS analysis of the reaction mixture employing a solution of 2,5-dihydroxybenzoic acid (DHB, 5 mg/mL in acetonitrile) as matrix showed the formation of fucosylated 5-aminopentyl **G0**. Control experiments lacking **HsFUT8** and GDP-Fuc were performed showing the unaltered 5-aminopentyl **G0** by MALDI-TOF MS (Supplementary Fig. 1).

Glycan array analysis. **HsFUT8** screening on glycan microarrays was performed as previously described with slight modifications¹³. Each subarray was incubated with a solution (200 μL) containing two different concentrations of **FUT8** (650 nM and 3.25 μM), GDP-Fuc (0.5 mM), MnCl₂ (10 mM) in Tris 25 mM, 150 mM NaCl pH = 7.5 at rt overnight. The supernatant was removed and the slide was washed with PBS and water. The introduction of Fuc residues by the action of **FUT8** was probed by incubation with *Aleuria aurantia* lectin (Vector laboratories, Burlingame, CA, USA) labeled with Alexa Fluor[™] 555 NHS succinimidyl ester (Thermo Fisher, AAL-555, 60 μg/mL in binding buffer: Tris 25 mM, 150 mM NaCl pH = 7.5 containing 0.01% Tween-20) at rt for 1 h. The microarray slide was washed with binding buffer, and water and fluorescence was analyzed on an Agilent G265BA microarray scanner system (Agilent Technologies, Santa Clara, USA). Quantification was performed with ProScanArray[®] Express (Perkin Elmer, Shelton, USA) and Microsoft Excel software. The average of mean RFU values after background subtraction and standard deviation for four replicate spots was represented as histograms employing GraphPad Prism 6 software.

Expression and purification of **HsFUT8.** The DNA sequence encoding amino acid residues 68–575 of the **HsFUT8** was codon optimized and synthesized by GenScript (USA) for expression in HEK293 cells (see Supplementary Table 2 for the codon optimized sequence). The DNA, containing at the 5'-end a recognition sequence for *KpnI*, and at the 3' end a stop codon and a recognition sequence for *XhoI*, was cloned into a modified pHLSec containing after the secretion signal sequence a 12xHis tag, a superfolder GFP³⁵ and a Tobacco Etch Virus (TEV) cleavage site, rendering the vector pHLSec-12His-GFP-TEV-**HsFUT8**. Both the synthesis of the **HsFUT8** construct and the engineered pHLSec together with the cloning of **HsFUT8** into pHLSec-12His-GFP-TEV were performed by GenScript.

pHLSec-12His-GFP-TEV-**HsFUT8** was transfected into HEK293F cell line (Thermo Fisher Scientific) as described below. Cells were grown in suspension in a humidified 37 °C and 8% CO₂ incubator with rotation at 125 r.p.m. Transfection was performed at a cell density of 2.5 × 10⁶ cell/mL in fresh media F17 serum-free media with 2% Glutamax and 0.1% P188. For each 150 mL of culture, 450 μg of the plasmid (1 μg/μL) was diluted to 135 μL with sterilized 1.5 M NaCl. This mixture was added to each 150 mL cell culture flask and incubated for 5 min in the incubator. After that, 1.35 mg of PEI-MAX (1 mg/mL) was mixed to 135 μL with sterilized 1.5 M NaCl and added to the cell culture flask. Cells were diluted 1:1 with pre-warmed media supplemented with valproic acid 24 h post-transfection to a final concentration of 2.2 mM. Cells were harvested 6 days post-transfection by

spinning down at 300 × *g* for 5 min, after which the supernatants were collected and centrifuged at 4,000 × *g* for 15 min.

Supernatant was dialyzed against buffer A (25 mM TRIS pH 7.5, 300 mM NaCl) and loaded into a His-Trap Column (GE Healthcare). Protein was eluted with an imidazol gradient in buffer A from 10 mM up to 500 mM. Buffer exchange to 25 mM TRIS pH 7.5, 150 mM NaCl (buffer B) was carried out using a HiPrep 26/10 Desalting Column (GE Healthcare). TEV protease was then added in a ratio 1:50 (TEV:protein) to the fusion construct in order to cleavage the His-GFP. After 20 h of reaction at 18 °C, the cleavage was satisfactorily verified through SDS-PAGE.

TEV protease and GFP were later removed from the solution using a His-Trap Column (GE Healthcare), and isolated **HsFUT8** was then loaded into a HiLoad 26/60 Superdex 75 Column (GE Healthcare), previously equilibrated with buffer B. Quantification of protein was carried out by absorbance at 280 nm using his theoretical extinction coefficient ($\epsilon_{280\text{ nm}}^{\text{HsFUT8}} = 90,760\text{ M}^{-1}\text{ cm}^{-1}$).

Crystallization and data collection. Crystals of the **HsFUT8** were grown by sitting drop experiments at 18 °C by mixing 0.5 μL of protein solution (4.5 mg/mL **HsFUT8**, 5 mM GDP and 5 mM **G0** in buffer B) with an equal volume of a reservoir solution (0.1 M carboxylic acids, 0.1 M buffer system 3 pH 8.5 and 30% precipitant mix 1 (Molecular Dimensions)). The crystals were cryoprotected in mother liquor containing 30% glycerol and flash frozen in liquid nitrogen.

Structure determination and refinement. Diffraction data were collected on the synchrotron beamline I24 of the Diamond Light Source (Harwell Science and Innovation Campus, Oxfordshire, UK) at a wavelength of 0.97 Å and a temperature of 100 K. Data were processed and scaled using XDS³⁶ and CCP4 (refs. 37,38) software packages. Relevant statistics are given in Table 1. The crystal structure was solved by molecular replacement with Phaser^{37,38} using the PDB entry 2DE0 as the template. Initial phases were further improved by cycles of manual model building in Coot³⁹ and refinement with REFMAC5 (ref. 40). Further rounds of Coot and refinement with REFMAC5 were performed to obtain the final structure. The final model was validated with PROCHECK; model statistics are given in Table 1. The AU of the C2 crystal contained two molecules of **HsFUT8**. The Ramachandran plots for the **HsFUT8**-GDP-**G0** show that 92.1%, 7.6%, 0.2% and 0% of the amino acids are in most favored, allowed, generously allowed and disallowed regions, respectively.

Molecular mechanics minimizations. GDP-Fuc was obtained from *CePoFUT1* in complex with GDP-Fuc (PDB entry: 3ZY6 (ref. 25)) and then it was superimposed on our crystal structure. The calculations were carried out on the complexes using AMBER 18 package, which was implemented with ffl45B, GLYCAM06 and GAFF force fields. The complex was immersed in a water box with a 10 Å buffer of TIP3P water molecules and neutralized by adding explicit Na⁺. A two-stage geometry optimization approach was performed with a total of 5000 minimization steps and using the default settings of AMBER 18. The first stage minimizes only the positions of solvent molecules and ions, and the second stage is an unrestrained minimization of all the atoms in the system.

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

Data availability

The crystal structure of the **HsFUT8**-GDP-**G0** complex was deposited at the RCSB PDB with accession code 6TKV. The source data underlying Fig. 2 are provided as a Source Data file. Other data are available from the corresponding author upon reasonable request.

Received: 5 January 2020; Accepted: 30 January 2020;

Published online: 20 February 2020

References

- Wang, X. et al. Dysregulation of TGF-beta1 receptor activation leads to abnormal lung development and emphysema-like phenotype in core fucose-deficient mice. *Proc. Natl. Acad. Sci. USA* **102**, 15791–15796 (2005).
- Schneider, M., Al-Shareffi, E. & Haltiwanger, R. S. Biological functions of fucose in mammals. *Glycobiology* **27**, 601–618 (2017).
- Wang, X. et al. Core fucosylation regulates epidermal growth factor receptor-mediated intracellular signaling. *J. Biol. Chem.* **281**, 2572–2577 (2006).
- Osumi, D. et al. Core fucosylation of E-cadherin enhances cell-cell adhesion in human colon carcinoma WiDr cells. *Cancer Sci.* **100**, 888–895 (2009).
- Zhao, Y. et al. Deletion of core fucosylation on alpha3beta1 integrin down-regulates its functions. *J. Biol. Chem.* **281**, 38343–38350 (2006).
- Nishima, W., Miyashita, N., Yamaguchi, Y., Sugita, Y. & Re, S. Effect of bisecting GlcNAc and core fucosylation on conformational properties of

- biantennary complex-type N-glycans in solution. *J. Phys. Chem. B* **116**, 8504–8512 (2012).
7. Andre, S. et al. Substitutions in the N-glycan core as regulators of biorecognition: the case of core-fucose and bisecting GlcNAc moieties. *Biochemistry* **46**, 6984–6995 (2007).
 8. Pereira, N. A., Chan, K. F., Lin, P. C. & Song, Z. The “less-is-more” in therapeutic antibodies: afucosylated anti-cancer antibodies with enhanced antibody-dependent cellular cytotoxicity. *MAbs* **10**, 693–711 (2018).
 9. Keeley, T. S., Yang, S. & Lau, E. The diverse contributions of fucose linkages in cancer. *Cancers (Basel)* **11**, 1241 (2019).
 10. Calderon, A. D. et al. Substrate specificity of FUT8 and chemoenzymatic synthesis of core-fucosylated asymmetric N-glycans. *Org. Biomol. Chem.* **14**, 4027–4031 (2016).
 11. Yang, Q., Zhang, R., Cai, H. & Wang, L. X. Revisiting the substrate specificity of mammalian alpha1,6-fucosyltransferase reveals that it catalyzes core fucosylation of N-glycans lacking alpha1,3-arm GlcNAc. *J. Biol. Chem.* **292**, 14796–14803 (2017).
 12. Ihara, H. et al. Crystal structure of mammalian alpha1,6-fucosyltransferase, FUT8. *Glycobiology* **17**, 455–466 (2007).
 13. Serna, S., Yan, S., Martin-Lomas, M., Wilson, I. B. & Reichardt, N. C. Fucosyltransferases as synthetic tools: glycan array based substrate selection and core fucosylation of synthetic N-glycans. *J. Am. Chem. Soc.* **133**, 16495–16502 (2011).
 14. Echeverria, B. et al. Chemoenzymatic synthesis of N-glycan positional isomers and evidence for branch selective binding by monoclonal antibodies and human C-type lectin receptors. *ACS Chem. Biol.* **13**, 2269–2279 (2018).
 15. Holm, L. & Laakso, L. M. Dali server update. *Nucleic Acids Res.* **44**, W351–W355 (2016).
 16. Urbanowicz, B. R. 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).
 17. Rocha, J. et al. Structure of *Arabidopsis thaliana* FUT1 reveals a variant of the GT-B class fold and provides insight into Xyloglucan fucosylation. *Plant Cell* **28**, 2352–2364 (2016).
 18. Valero-Gonzalez, J. et al. A proactive role of water molecules in acceptor recognition by protein O-fucosyltransferase 2. *Nat. Chem. Biol.* **12**, 240–246 (2016).
 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).
 20. Chen, C. I. et al. Structure of human POFUT2: insights into thrombospondin type 1 repeat fold and O-fucosylation. *EMBO J.* **31**, 3183–3197 (2012).
 21. Corzana, F. et al. The pattern of distribution of amino groups modulates the structure and dynamics of natural aminoglycosides: implications for RNA recognition. *J. Am. Chem. Soc.* **129**, 2849–2865 (2007).
 22. Canales, A. et al. Breaking pseudo-symmetry in multiantennary complex N-glycans using lanthanide-binding tags and NMR pseudo-contact shifts. *Angew. Chem. Int. Ed. Engl.* **52**, 13789–13793 (2013).
 23. Kadirvelraj, R. et al. Human N-acetylglucosaminyltransferase II substrate recognition uses a modular architecture that includes a convergent exosite. *Proc. Natl Acad. Sci. USA* **115**, 4637–4642 (2018).
 24. Moremen, K. W. & Haltiwanger, R. S. Emerging structural insights into glycosyltransferase-mediated synthesis of glycans. *Nat. Chem. Biol.* **15**, 853–864 (2019).
 25. Lira-Navarrete, E. et al. Structural insights into the mechanism of protein O-fucosylation. *PLoS ONE* **6**, e25365 (2011).
 26. Kotzler, M. P. et al. Donor assists acceptor binding and catalysis of human alpha1,6-fucosyltransferase. *ACS Chem. Biol.* **8**, 1830–1840 (2013).
 27. Tseng, T. H. et al. Substrate preference and interplay of fucosyltransferase 8 and N-acetylglucosaminyltransferases. *J. Am. Chem. Soc.* **139**, 9431–9434 (2017).
 28. Lira-Navarrete, E. et al. Dynamic interplay between catalytic and lectin domains of GalNAc-transferases modulates protein O-glycosylation. *Nat. Commun.* **6**, 6937 (2015).
 29. Albesa-Jove, D. & Guerin, M. E. The conformational plasticity of glycosyltransferases. *Curr. Opin. Struct. Biol.* **40**, 23–32 (2016).
 30. Kurochkina, N. & Guha, U. SH3 domains: modules of protein-protein interactions. *Biophys. Rev.* **5**, 29–39 (2013).
 31. Sun, B. et al. A simplified procedure for gram-scale production of sialylglycopeptide (SGP) from egg yolks and subsequent semi-synthesis of Man3GlcNAc oxazoline. *Carbohydr. Res.* **396**, 62–69 (2014).
 32. Varki, A. & Diaz, S. The release and purification of sialic acids from glycoconjugates: methods to minimize the loss and migration of O-acetyl groups. *Anal. Biochem.* **137**, 236–247 (1984).
 33. Seko, A. et al. Occurrence of a sialylglycopeptide and free sialylglycans in hen’s egg yolk. *Biochim. Biophys. Acta* **1335**, 23–32 (1997).
 34. Serna, S., Etxebarria, J., Ruiz, N., Martin-Lomas, M. & Reichardt, N. C. Construction of N-glycan microarrays by using modular synthesis and on-chip nanoscale enzymatic glycosylation. *Chemistry* **16**, 13163–13175 (2010).
 35. Moremen, K. W. et al. Expression system for structural and functional studies of human glycosylation enzymes. *Nat. Chem. Biol.* **14**, 156–162 (2018).
 36. Kabsch, W. Xds. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132 (2010).
 37. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 235–242 (2011).
 38. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).
 39. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
 40. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).

Acknowledgements

We thank the Diamond Light Source (Oxford, UK) synchrotron beamline I24 (experiment number mx20229-19). We thank ARAID, MEC (CTQ2013-44367-C2-2-P, BFU2016-75633-P, CTQ2017-90039-R, RTC-2017-6126-1 and RTI2018-099592-B-C21), Maria de Maeztu Units of Excellence Program from the Spanish State Research Agency (MDM-2017-0720) and Gobierno de Aragón (E34_R17 and LMP58_18) with FEDER (2014-2020) funds for ‘Building Europe from Aragón’ for financial support. The research leading to these results has also received funding from the FP7 (2007-2013) under BioStruct-X (grant agreement No. 283570 and BIOSTRUCTX_5186).

Author contributions

R.H.-G. designed the crystallization construct and solved the crystal structure. A.G.-G. and L.C.-L. purified the enzyme and crystallized the complex. R.H.-G. refined the crystal structure. F.C. performed the molecular mechanics calculations. S.S., R.A. and N.C.R. performed the synthesis of G0 and the mass spectrometry experiments together with the microarray of synthetic N-glycans experiment. R.H.-G. wrote the article with mainly the contribution of N.C.R. and F.C. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-14794-z>.

Correspondence and requests for materials should be addressed to R.H.-G.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020