Supporting Information

Characterizing human α-1,6-fucosyltransferase (FUT8) substrate specificity and structural similarities with related fucosyltransferases

Bhargavi M. Boruah[1,†], Renuka Kadirvelraj[2,†], Lin Liu[1], Annapoorani Ramiah[1], Chao Li[3], Guanghui Zong[3], Gerlof P. Bosman[4], Jeong-Yeh Yang[1], Lai-Xi Wang[3], Geert-Jan Boons[1], Zachary A. Wood[2*], and Kelley W. Moremen[1,2*]

[1]Complex Carbohydrate Research Center, University of Georgia, Athens, GA;
[2]Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA;
[3]Department of Chemistry and Biochemistry, University of Maryland, College Park, MD;
[4]Department of Chemical Biology and Drug Discovery, Utrecht Institute for Pharmaceutical Sciences, and Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, The Netherlands.

[†] These authors contributed equally to this work.

[*] Correspondence: Kelley Moremen: email: moremen@uga.edu and Zachary A. Wood: email: zaw@uga.edu

This Supplementary Information includes:

Supplementary Table S1: Kinetic parameters for wild type human FUT8 using GDP-Fucose as donor and various glycan acceptor substrates.
Supplementary Table S2. Data collection and refinement statistics
Supplementary Table S3: Kinetic parameters for wild type human FUT8 and mutants.
Supplementary Table S4: Primer sequences used for site directed mutagenesis.
Supplementary Figure S1. FUT8 expression, purification, and tag/glycan cleavage.
Supplementary Figure S2. Structural alignment of fucosyltransferase structures using the PROMALS3D server.
Supplementary Figure S3. Sequence and structural alignment of FUT8 with other related GTs.

Supplementary Table S1: Kinetic parameters for wild type human FUT8 using GDP-Fucose as donor and various glycan acceptor substrates.

Acceptor Substrates[a]	k_{cat}	K_m	k_{cat}/K_m
	s^{-1}	μM	$mM^{-1}s^{-1}$
M5N2-Asn-Fmoc	0.2 ± 0.1	949 ± 506	0.21 ± 0.22
NM5N2-Asn	6.9 ± 0.8	214 ± 64	32 ± 13
M3N2-Asn	N.D.		
A1-Asn	5.9 ± 0.7	22.5 ± 7.7	260 ± 120
A2-Asn	13.7 ± 0.8	51.7± 9.7	265 ± 65
G1-Asn	8.7 ± 0.6	159 ± 30	55 ± 14
A3-Asn	7.0 ± 0.5	556 ± 82	12 ± 3
A3'-Asn	9.6 ± 0.9	149 ± 39	64 ± 23
A4-Asn	N.D.		

N.D. not detectable

[a] Fucosyltransferase assays were performed in duplicate with the respective glycan substrates using a GDP-Glo assay format as described in Experimental Procedures. Steady state parameters of k_{cat} and K_m were calculated using a GDP standard curve and nonlinear curve fitting in GraphPad Prism 6 software.

Supplementary Table S2. Data collection and refinement statistics

Data collection	**Fut8:GDP**	**Fut8:GDP:A2**	**Fut8:GDP:A3'**	**Fut8:GDP:A3**	**Fut8:GDP:NM5N2**
Wavelength (Å)	1.0000	1.0000	1.0000	1.0000	1.0000
Space group	$P6_5$	$P6_5$ 2 2	$P6_5$ 2 2	$P6_5$	$P6_5$
Cell dimensions (a, b, c in Å)	173.5, 173.5 207.4	149.9, 149.9, 480.9	150.5, 150.5, 480.3	151.7, 151.7, 474.1	154.0, 154.0, 466.9
Molecules in asymmetric unit	4	2	2	8	8
Completeness (%)	99.2 (97.1)[a]	99.3 (97.1)[a]	89.4 (93.6)[a]	99.8 (100)[a]	99.3 (99.9)[a]
Total number of reflections	837360 (49014)	1290703 (103944)	149167 (11496)	1120885 (87009)	367570 (28430)
Unique reflections	165616 (11970)	124545 (8849)	44367 (3352)	219502 (16316)	102079 (7647)
Redundancy	5.1 (4.1)	10.4 (11.7)	3.4 (3.4)	5.1 (5.3)	3.6 (3.7)
$I / \sigma(I)$	10.4 (1.5)	18.4 (2.3)	7.9 (1.8)	10.7 (1.4)	5.7 (1.2)
R_{meas}[b] (%)	15.8 (109.7)	15.3 (146.6)	19.2 (92.9)	12.0 (105.3)	25.9 (139.7)
$CC_{1/2}$ (%)	99.5 (55.6)	99.9 (76.1)	99.0 (68.9)	99.7 (50.5)	98.5 (50.9)
Refinement					
Resolution and highest-resolution shell (Å)	2.25 (2.31-2.25)	2.4 (2.46-2.4)	3.3 (3.39-3.3)	2.47 (2.53-2.47)	3.2 (3.28-3.2)
R_{work} / R_{free}	0.17 / 0.20	0.205 / 0.23	0.20 / 0.22	0.167 / 0.20	0.19 / 0.236
No. of atoms Protein / Ligand / Solvent	14987 / 98 / 1351	7639 / 236 / 634	7600 / 268 / 26	30384 / 474 / 769	30432 / 752 / 264
B-factor ($Å^2$) Protein / Ligand / Solvent	41.8 / 32.4 / 50.1	50.1 / 45.8 / 56.1	69 / 66 / 73.6	50.8 / 50.3 / 54.4	61.8 / 68.7 / 53.7
Wilson *B*-factor ($Å^2$)	31.7	43	77.1	54.6	64.1
Stereochemical Ideality					
Bond lengths ($Å^2$)	0.009	0.010	0.01	0.005	0.006
Bond angles (°)	0.99	1.26	1.3	0.97	0.98
Ramachandran favored (%)	97.2	97.2	95.6	96.9	97.0
Ramachandran allowed (%)	2.8	2.8	4.4	3.1	3.0
PDB code	6X5H	6X5R	6X5S	6X5T	6X5U

[a]Values in parentheses are for highest-resolution shell as defined under Resolution.

[b]R_{meas} is the redundancy independent merging *R*-factor of Diederichs and Karplus (118).

Supplementary Table S3: Kinetic parameters for wild type human FUT8 and mutants using GDP-Fucose as donor and A2 substrate as acceptor[a].

Enzyme form	GFP fluorescence[b]	Acceptor				Donor			
		k_{cat}	K_m	k_{cat}/K_m	$\frac{k_{cat}/K_m}{k_{cat}/K_m (wt)}$ [d]	k_{cat}	K_m	k_{cat}/K_m	$\frac{k_{cat}/K_m}{k_{cat}/K_m (wt)}$ [d]
		s^{-1}	μM	$\mu M^{-1}s^{-1}$		s^{1}	μM	$\mu M^{-1}s^{-1}$	
WT(293F)[e]	624	13.7 ± 0.8	51.7 ± 9.7	0.26		11.2 ± 0.8	67.9 ± 13.0	0.16	
WT(293S)[c]	434	5.2 ± 0.4	57.6 ± 18.0	0.09		2.0 ± 0.3	71.6 ± 22.8	0.03	
K216A[e]	170	0.08 ± 0.01	53.8± 16.2	0.0015	0.006	0.05 ± 0.01	66.0 ± 32.5	0.0007	0.004
D295A[e]	381	N.D.				N.D.			
Q470A[e]	547	0.23 ± 0.03	85.9 ± 26.8	0.0027	0.010	0.25 ± 0.19	352 ± 381	0.0007	0.004
R473A[e]	161	N.D.				N.D.			
D494A[e]	165	0.06 ± 0.01	55.5 ± 15.7	0.001	0.0045	0.04 ± 0.01	227 ± 129	0.0002	0.001
D495A[e]	388	0.04 ± 0.01	152 ± 98	0.0003	0.0010	0.03 ± 0.01	271 ± 173	0.0001	0.0006
Q502A[e]	195	0.16 ± 0.02	13.1 ± 5.7	0.01	0.05	0.016 ± 0.003	79.8 ± 35.4	0.0002	0.001
H535A[e]	311	0.70 ± 0.03	30.4 ± 4.7	0.02	0.09	-	> 1000[f]	-	-
K541A[e]	428	5.8 ± 0.4	38.8 ± 9.8	0.15	0.58	5.2 ± 0.6	183± 35	0.03	0.18
Y498A[e]	301	N.D.				N.D.			
R365A[e]	357	N.D.				N.D.			
D368A[e]	271	N.D.				N.D.			
K369A[e]	71	N.D.				N.D.			
E373A[e]	363	N.D.				N.D.			
Y382A[e]	197	N.D.				N.D.			
D409A[e]	253	N.D.				N.D.			
D453A[e]	191	N.D.				N.D.			
S469A[e]	285	N.D.				N.D.			

[a] Fucosyltransferase assays were performed in duplicate with the A2-Asn acceptor and GDP Fucose as donor substrates using a GDP-Glo assay format as described in Experimental Procedures.

[b] The relative expression and secretion of the GFP-FUT8 fusion proteins in transiently transfected HEK293 cells was determined by measuring the fluorescence of the recombinant proteins secreted into the media.

[c] The GFP-FUT8 fusion protein expressed in HEK293S (GnTI-) cells was purified by Ni^{2+}-NTA chromatography, cleaved to remove tag sequences and further purified as described for crystallography of the enzyme catalytic domain.

[d] Values for k_{cat}/K_m for all FUT8 mutants were compared with k_{cat}/K_m values for wild type FUT8 expressed in HEK293F (wild type) cells.

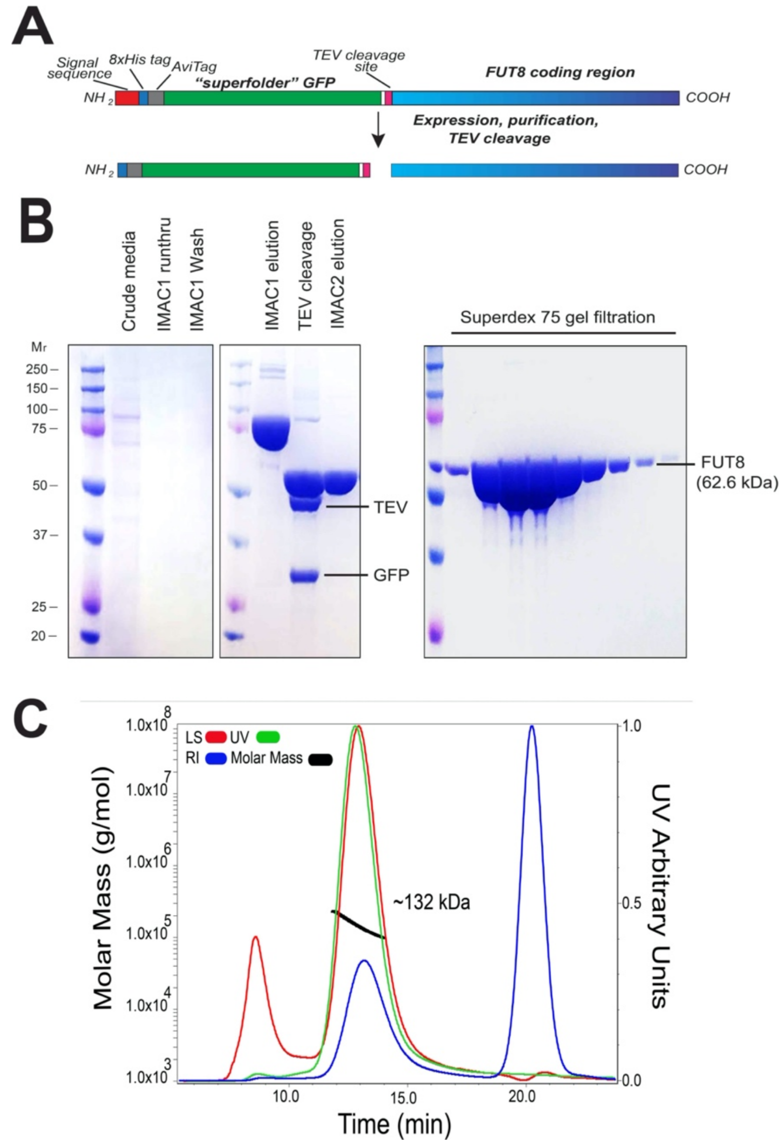
[e] The GFP-FUT8 fusion protein (and all FUT8 mutants) were expressed in HEK293F (wild type) cells and purified by Ni^{2+}-NTA chromatography. The fusion protein tags were retained for the indicated wild type FUT8 and mutant forms of the enzyme during kinetic analyses.

[f] Catalytic activity was detected, but values for k_{cat} and K_m could not be determined.

N.D. not detectable

Supplementary Table S4: Primer sequences used for site directed mutagenesis. Lower case letters represent bases that were mutated using the Q5 site-directed mutagenesis kit (see Experimental Procedures).

Mutant	Primer Name	Primer sequence
K216A	Mut-1_F	TAATATCAACgcaGGCTGTGGCTATG
	Mut-1_R	CACACCAGCTTTTTGGCT
D295A	Mut-2_F	TCCCATTGTAgccAGTCTTCATC
	Mut-2_R	AGCTCGACCACTTGAACA
Q470A	Mut-3_F	TTTTTCATCCgcgGTCTGTCGAGTTG
	Mut-3_R	GTACACACTAGGAAGTCTG
R473A	Mut4_F	CCAGGTCTGTgcaGTTGCTTATG
	Mut4_R	GATGAAAAAGTACACACTAG
D494A	Mut5_F	CCATTCTTTAgctGACATCTACTATTTTG
	Mut5_R	AAGTTTGCAGAGGCATCAG
D495A	Mut6_F	TTCTTTAGATgccATCTACTATTTTGGG
	Mut6_R	TGGAAGTTTGCAGAGGCA
Q502A	Mut7_F	TTTTGGGGGCgcgAATGCCCACAATC
	Mut7_R	TAGTAGATGTCATCTAAAGAATGG
H535A	Mut8_F	GGCTGGAAATgctTGGGATGGCTATTC
	Mut8_R	ACACCAATGATATCTCCAG
K541A	Mut9_F	TGGCTATTCTgcaGGTGTCAACAG
	Mut9_R	TCCCAATGATTTCCAGCC
Y498A	Mut10_F	TGACATCTACgcTTTTGGGGGCC
	Mut10_R	TCTAAAGAATGGAAGTTTGC
R365A	Mut11_F	AGTCCATGTCgcaCGCACAGAC
	Mut11_R	CCAATAACTGGATGTTTG
D368A	Mut12_F	AGACGCACAgcTAAAGTGGGAAC
	Mut12_R	GACATGGACTCCAATAAC
K369A	Mut13_F	ACGCACAGACgcaGTGGGAACAG
	Mut13_R	CTGACATGGACTCCAATAAC
E373A	Mut14_F	GTGGGAACAgcaGCTGCCTTC
	Mut14_R	TTTGTCTGTGCGTCTGAC
Y382A	Mut15_F	CATTGAAGAGgccATGGTGCATGTTGAAG
	Mut15_R	GGATGGAAGGCAGCTTCT
D409A	Mut16_F	TTGGCCACAGcTGACCCTTCT
	Mut16_R	ATACACTCTTTTTTTGTCCACTTG
D453A	Mut17_F	GTGATCCTGGcTATACATTTTCTC
	Mut17_R	TCCACGAAGTGAATTTTC
S469A	Mut18_F	TACTTTTTCAgccCAGGTCTG
	Mut18_R	CACACTAGGAAGTCTGCC

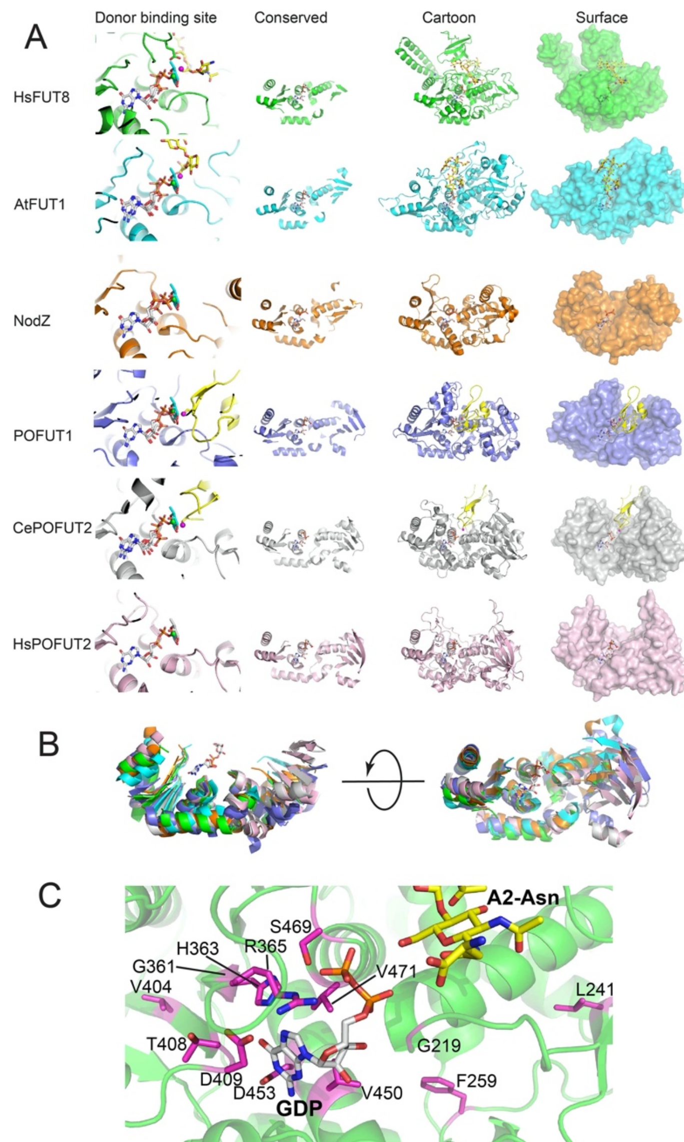


Supplementary Figure S1. FUT8 expression, purification, and tag/glycan cleavage. (**A**) Diagrammatic representation of the recombinant FUT8 fusion protein coding region is shown. The fusion protein has an NH_2-terminal signal sequence followed by an 8xHis tag, AviTag, superfolder GFP, TEV protease cleavage site, and the catalytic domain of FUT8. (**B**) Expression of the recombinant product in HEK293F 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 resulted in removal of the tag sequences (*TEV cleavage*). Ni^{2+}-NTA chromatography separated the unbound FUT8 catalytic domain (*IMAC2 elution*) from the bound tag sequences and TEV protease, as the latter were both His-tagged. The enzyme was further purified over Superdex 75 (*Superdex 75 gel filtration*). (**C**) The purified FUT8 catalytic domain following cleavage with TEV was further characterized by size exclusion-multiangle light scattering (SEC-MALS, see Materials and Methods). 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 (132 kDa) is in close agreement with a dimer of the calculated mass for the FUT8 catalytic domain following cleavage with TEV protease (62.6 kDa)

```
Conservation:
NodZ_3six.pdb             -----------------------------------------------------------------------------------------
MmPOFUT1_5ky3.pdb         -----------------------------------------------------------------------------------------
HsFUT8:GDP:A2Asn       1  ------LGK--DHEILRRRIENGAKELWFFLQSELKKLKNLEGNELQRHADEFLLDLGHHERSIMTDLYYLSQTDGAGDWREKEAKDLTELVQRRITYLQ   92
HsPOFUT2_4ap6.pdb         -----------------------------------------------------------------------------------------
CePOFUT2_5foe.pdb         -----------------------------------------------------------------------------------------
AtFut1_5kor.pdb        1  INSDKLLGGLLASGFDEDSCLSRYQSVHYRKP--------------------------------------------SPYKPSSYLISKLRNYEKLH   52
Consensus_aa:             ..........................................................................................
Consensus_ss:

Conservation:                                      6           7  6                     6              9                           7   6
NodZ_3six.pdb          1  -----------------------KERFVISRRR--TGFGDCLWSLASAWSYAQRTG-----RTLVIDWRGS-----CYVEQPFSNAFPAFFEPVE-DIA   63
MmPOFUT1_5ky3.pdb      1  ------PSWD--------------LAGYLLYCPC-MGRFGNQADHFLGSLAFAKLLN-----RTLAVPPWIE-YQHHKPPFTN-LHVSYQKYFKLEPLQA   72
HsFUT8:GDP:A2Asn      93  NPK----DCS--------------KAKKLVCNINKGCGYGCQLHHVVYCFMIAYGTQ-----RTLILESQNW-----RYA-----TGGWETVFRP-----   154
HsPOFUT2_4ap6.pdb      1  ---------S--------------RRRYLLYDVNPPAGFNLRRDVYIRIASLLKTLL-KTEEWVLVLPPWGRLYHWQSPDIHQ-VRIPWSEFFDLPSLNK   75
CePOFUT2_5foe.pdb      1  -----------------------EKKFLLYDVNFGEGFNLHRRDVYMRVANTVRSLRDSGENYILVIPPWGRE-----------VALSWRLFFDLESLNF   65
AtFut1_5kor.pdb       53  KRCGPGTESYKKALKQLDQEHIDGECKYVVWISF--SGLGNRILSLASVFLYALLTD-----RVLLVDRG----------------KDMDDLFCEP---F   126
Consensus_aa:             .......................c.+@llh....sG@s.p...hh.hh.hhb.h......hLlls................p.shp.hF......
Consensus_ss:                                     eeeee                hhhhhhhhhhhhhhhh        eeeee                     hhhh   hh  hhh

Conservation:                                                                         6
NodZ_3six.pdb         64  GVPVICDDRVNQ---------LSF-----EGPFFP------------------RWW-----NRPSIDCI---NRP------DEQIFR-----------ER   106
MmPOFUT1_5ky3.pdb     73  YHRVVSLEDFMENLAPSHWPPEK-----RVAYCF-----------------EVAAQRSPDKKTCPMKE---GNPFGPFWDQFHVSFNKSEL-------T   140
HsFUT8:GDP:A2Asn     155  VSETCTDR--------SGIS---------TGHWS--------------------------------------------------------------------   171
HsPOFUT2_4ap6.pdb     76  NIPVIEYEQFIA---ESGGPFI------DQVYVLQSYAEGWKEGTWEEKVDERPCIDQ-----LLYSQDKHEYYRGWFWGYEETRGLNVSC------LS   154
CePOFUT2_5foe.pdb     66  FIPVIEFEDFLD----ENRPI-------DQVIYLQHYAEGW-GTEYVRKFEKRSCLPP---AESHYKQVEEFKWKGWFYSYEDVYSRNFQC------VS   143
AtFut1_5kor.pdb      127  LGMSWLL--PLD----FPMTDQFDGLNQESSRC--------------YGYMVKN-----------------------QVITI-SHLYLHLVHDYGDHD   180
Consensus_aa:             ...hh..c...p.....s.........shh.........................................................bpl..........p
Consensus_ss:                eehhhhhh      h          eeee                                                      eeeeeee          e

Conservation:                                       6
NodZ_3six.pdb        107  DELTEIFQAREDSEA-----------NTIVCDA------CLMWRCS-----------------------EEAERLIFRNIKIRSEIRARIDALYEEHFSG----   166
MmPOFUT1_5ky3.pdb    141  GISF-SASYKEQWTQRFPA---KEHPVLALPGAP----AQFPVLEEH----------------------RELQKY---MVWSDE-MVRTGEALISAHLV-----   205
HsFUT8:GDP:A2Asn     172  ----------GEV-KDKN-----VQVVELPIVDSLHPRPPYLPLA-VPEDLADRLVRVHGDPAVWWVSQFVKYLIRPQPWLEKEIEEATKKLGFK----   249
HsPOFUT2_4ap6.pdb    155  VQG--SASI--VAP-LLLRN--TSARSVMLDRAENL---LHDHYGG-----------------------KEYWDTRRS-MVKARHLREVGDEFRSRHLNSTDDA   224
CePOFUT2_5foe.pdb    144  IQG--DSGT--LKD-LIKHSNFSESTSIMVDRAETI---LHEHYGE-----------------------VDYWKARRS-MRYSNDLVDVADAFRKKYLDSDDKR   215
AtFut1_5kor.pdb      181  KMFF--CEG--DQT-FIGK-----VPWLIVKTDNYF--VPSLWLIPGFDDELNKLFP----QKA-TVFHHLGRYLFHPTNQVWGLVTRYYEAYLSH----   259
Consensus_aa:             .b...hp.....s.........h.slhls.h.........h..........................h.ch....h.pp.l...-.h.p.@hs.....
Consensus_ss:                 hhh   hhh  hh        eeee                                          hhhhhhhhhhh    hhhhhhhhhhhhhhhh

Conservation:                                   967 9  6
NodZ_3six.pdb        167  ---------------------HSIIGVHVRHAD-----------------------------------SELALHQVCMAIR--KAKALSY-----P   199
MmPOFUT1_5ky3.pdb    206  ---------------------RPYVGIHLRIGSDWKNACAMLKDGTAGSHFMASPQCVGYSRSTATPLTMTMCLPDLKEIQRAVT--LWVRALN------   276
HsFUT8:GDP:A2Asn     250  ---------------------HPVIGVHVRRTDKVGTEA------------------------------AFHPIEEYMVHVE--EHFQLL-ARRMQV   292
HsPOFUT2_4ap6.pdb    225  DRIPFQEDWMKMKVKLGSALGGPYLGVHLRRKDFIWGHR-------------------------------QDVPSLEGAVRKIR--SLMKTHR------   284
CePOFUT2_5foe.pdb    216  DKTKLVDDWTKEKPR-RTAIGGPYLGIHWRRRDFIYAKK-------------------------------AQLPTIPGTAKILQ--DLCKKLD------   274
AtFut1_5kor.pdb      260  ---------------------ADEKIGIQVRVF--DEDPG-------------------------------PFQHVMDQISSCTQKEKLLPEVDTLVNTP   305
Consensus_aa:             .....................shlGlHlR..s..s.....................................b.slc.h..hlp..phh..h......
Consensus_ss:                                    eeeeeeee   hhhh                                   hh  hhhhhhhhh   hhhhh
                                               Motif 1

Conservation:                 7   77                                                     7  7    6       6      9 7
NodZ_3six.pdb        200  KPVKVFLCTDS-AQVLDQVSGLF----------PDVFAVPK-----RSAE-MGIEGGASALIDMYLLARCATVIRFPPTSAFTRYARLLV-------PR   274
MmPOFUT1_5ky3.pdb    277  -ARSVYIATDS-ESYVSEIQQLFKD-------KVRVVSLKP----------------EVAQIDLYILGQADHFIG-NCVSSFTAFVKRERDLHG----RQ   346
HsFUT8:GDP:A2Asn     293  DKKRVYLATDD-PSLLKEAKTKYPN-------YEFISDNSISWSAGLHNR-YTENSLRGVILDIHFLSQADFLVC-TFSSQVCRVAYEIMQTLHPDASAN   382
HsPOFUT2_4ap6.pdb    285  -LDKVFVATDAVRKEYEELKKLLP--------EMVRFEP-----TWEELELYKDGGVAIIDQWICAHARFFIG-TSVSTFSFRIHEREILGLDPKTT   367
CePOFUT2_5foe.pdb    275  -LQKIYLATDAPDQEVDELKALLNGE-------LEVYRFTD---------TQKLNDGQIAIIDQYLCAHAAYFIG-SYESTFTFRIQEDREIIGFPISTT   356
AtFut1_5kor.pdb      306  KHKAVLVTSLN-AGYAENLKSMYWEYPTSTGEIIGVHQPSQ---------EGYKMHNGKALAEMYLLSLTDNLVT-SAWSTFGYVAQGLGG-------LK   387
Consensus_aa:             ..pbV@ltTDs..pbhcplp.hh.........lhp.p............b....AllDb@lhtptshhls.s..SsFs..hp.b.........p
Consensus_ss:                eeeee      hhhhhhhhhh          eeeee                      hhhhhhhhh     eeee      eehhhhhhhh
                      Motif 2                                                                          Motif 3

Conservation:                 6
NodZ_3six.pdb        275  IIEFDL------------------------------------SNPGHLT--MIDNP---------------------------------------------   292
MmPOFUT1_5ky3.pdb    347  SSFFGMD-------------------------------------------------------------------------------------------   353
HsFUT8:GDP:A2Asn     383  FHSLDDIYYFGGQNAHNQIAIYAHQPRTADEIPMEPGDIIGVAGNHWDGYSKGVNRKLGRTGLYPSYKVREKIE----------TVK---YPTYPE--   465
HsPOFUT2_4ap6.pdb    368  YNRFCGD---------------------------------QEKACE--QPTH-WKITY-----------------------------------------   389
CePOFUT2_5foe.pdb    357  FNRLCPD---------------------------------TEPTCE--QPAK-WKIVYSS--------------------------------------   380
AtFut1_5kor.pdb      388  PWILYRPENRTT----------------------------PDPSCG--RAMS-MEPCFHSPPFYDCKAKTGIDTGTLVPHVRHCEDISWGLKLV   450
Consensus_aa:             ...hs.....................................s..p.......p...........................................
Consensus_ss:
```

Supplementary Figure S2. Structural alignment of fucosyltransferase structures using the PROMALS3D server. The structure of the human FUT8:GDP:A2-Asn complex was aligned with the mouse POFUT1:GDP-Fuc:EGF domain complex (PDB 5KY3 (38)), the *C. elegans* POFUT2:GDP:thrombospondin repeat complex (PDB 5FOE (39), the human POFUT2:GDP-Fuc complex (PDB 4AP6 (40)), the NodZ:GDP complex (PDB 3SIX (35)), and the AtFUT1:GDP:xylo-oligosaccahride complex (PDB 5KOR (29) using default parameters in the PROMALS3D server (97). 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 three conserved sequence elements previously identified by primary sequence alignment (94, 95). Conserved secondary structure elements for the fucosyltransferases that are indicated in the structural alignment in ***Fig. 10*** are indicated by the blue boxes. The putative catalytic bases for FUT8 and *C. elegans* POFUT1 are indicated with the bold magenta 'E' with a green circle.



Supplementary Figure S3. Sequence and structural alignment of FUT8 with other related GTs. The structures of FUT8:GDP:A2-Asn, mouse POFUT1:GDP-Fuc:EGF domain complex (PDB 5KY3 (38)), *C. elegans* POFUT2:GDP:thrombospondin repeat complex (PDB 5FOE (39), human POFUT2:GDP-Fuc complex (PDB 4AP6 (40)), NodZ:GDP complex (PDB 3SIX (35)), and *Arabidopsis* FUT1:GDP:xylo-oligosaccahride complex (PDB 5KOR (29) were aligned in Coot (111) followed by manual adjustment in PyMOL. **(A)** Individual structures are displayed in surface representation (*Surface*) or cartoon representation (*Cartoon*) with bound GDP or GDP-Fuc shown in white stick representation. Acceptor ligands are shown as yellow sticks (glycan acceptors) or cartoon (POFUTs) representations. Conserved regions identified in the PROMALS3D alignment (blue boxes in *Fig. 9*) were extracted and shown as cartoon representations (*Conserved*). A zoom-in to the GDP-Fuc donor binding site (*Donor Binding Site*) is shown with the respective protein in cartoon representation and the donor in stick representation. For each of the donor binding site structures a representation of GDP-Fuc derived from the human POFUT2 structure is also superimposed to indicate the position of the bound intact sugar donor where the other complexes contained only GDP. The latter GDP-Fuc molecule is represented with the Fuc residue in cyan sticks and the C1 as a green sphere. When the acceptor structure is present, the hydroxyl nucleophile is represented as a magenta sphere. (**B**) The conserved structural elements identified in the PROMALS3D analysis (*Conserved* column in *Panel A*) are displayed as an overlay of the respective structures in two

different orientations. (**C**) Thirteen residues were completely conserved among the six structures. These conserved residues are shown within the FUT8:GDP:A2-Asn structure (magenta stick representation), where 9 residues are associated with the GDP binding site. GDP shown in white stick representation and the A2-Asn acceptor shown in yellow stick representation.