

Structure and Mechanism of *Helicobacter pylori* Fucosyltransferase

A BASIS FOR LIPOPOLYSACCHARIDE VARIATION AND INHIBITOR DESIGN^{*[5]}

Received for publication, November 3, 2006, and in revised form, January 12, 2007. Published, JBC Papers in Press, January 24, 2007, DOI 10.1074/jbc.M610285200

Han-Yu Sun^{†§1}, Sheng-Wei Lin^{†§1}, Tzu-Ping Ko^{§1}, Jia-Fu Pan[§], Chia-Ling Liu^{†§}, Chun-Nan Lin[§], Andrew H.-J. Wang^{†§¶||2}, and Chun-Hung Lin^{†§¶||3}

From the [†]Institute of Biochemical Sciences, National Taiwan University, Taipei 10642 and the [§]Institute of Biological Chemistry, the [¶]Genomics Research Center, and the ^{||}National Core Facility of High-Throughput Protein Crystallography, Academia Sinica, No.128 Academia Road Section 2, Nan-Kang, Taipei 11529, Taiwan

Helicobacter pylori α 1,3-fucosyltransferase (FucT) is involved in catalysis to produce the Lewis x trisaccharide, the major component of the bacteria's lipopolysaccharides, which has been suggested to mimic the surface sugars in gastric epithelium to escape host immune surveillance. We report here three x-ray crystal structures of FucT, including the FucT-GDP-fucose and FucT-GDP complexes. The protein structure is typical of the glycosyltransferase-B family despite little sequence homology. We identified a number of catalytically important residues, including Glu-95, which serves as the general base, and Glu-249, which stabilizes the developing oxonium ion during catalysis. The residues Arg-195, Tyr-246, Glu-249, and Lys-250 serve to interact with the donor substrate, GDP-fucose. Variations in the protein and ligand conformations, as well as a possible FucT dimer, were also observed. We propose a catalytic mechanism and a model of polysaccharide binding not only to explain the observed variations in *H. pylori* lipopolysaccharides, but also to facilitate the development of potent inhibitors.

Helicobacter pylori is a serious human pathogen that causes gastritis and both peptic and duodenal ulcers (1, 2). The pathogen is associated with an increased risk for development of both gastric adenocarcinoma (3) and mucosa-associated lymphoma (4). *H. pylori* is highly adapted to colonize in human gastric mucosa, where it may remain for decades or even a lifetime. The specific infection in the stomach is caused by the attachment of adhesion proteins, such as BabA and SabA, to specific glycoconjugates on the gastric epithelial cell surface (5, 6).

* This work was supported by Academia Sinica and the National Science Council, Taiwan (Grant NSC-94-3112-B-001-010-Y to A. H.-J. W. and Grant NSC-94-2113-M-001-0028 to C. H. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (codes 2NZW, 2NZX, and 2NZY) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

[5] The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Table S1 and Figs. S1–S3.

¹ These authors contributed equally to this work.

² To whom correspondence may be addressed. Tel.: 886-2-2788-1981; Fax: 886-2-2788-2043; E-mail: ahjwang@gate.sinica.edu.tw.

³ To whom correspondence may be addressed. Tel.: 886-2-2789-0110; Fax: 886-2-2651-4705; E-mail: chunhung@gate.sinica.edu.tw.

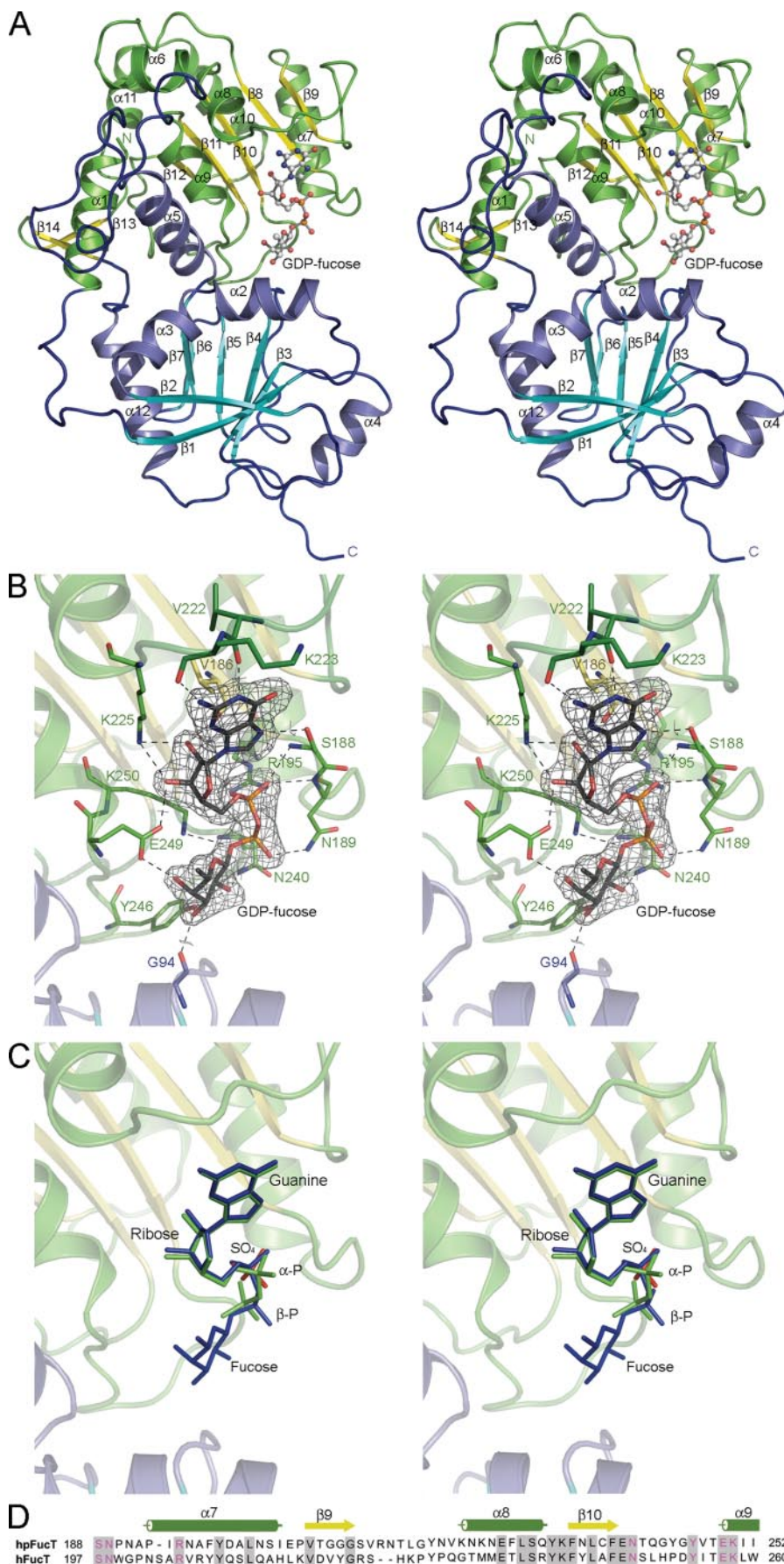
Although *H. pylori* elicits local as well as systemic antibody response (7), location at such a specific niche permits it to escape elimination by the host immune response.

The lipopolysaccharides (LPSs)⁴ of *H. pylori* contain fucosylated oligosaccharides, predominantly the type II blood group antigens Lewis x and Lewis y (8) in addition to the minor type I Lewis a and Lewis b (9). *H. pylori* fucosyltransferases (FucTs) are responsible for the last steps in the biosynthesis of Lewis antigens. The molecular mimicry of host cell surface antigens has been suggested to mask the pathogen from host immune surveillance and thus plays an important role in colonization and long term infection in the stomach (10). Furthermore, *H. pylori* continuously alters the expression of Lewis antigens, a process known as phase variation, to generate several LPS variants in one bacterial population and to display structural heterogeneity. The dynamic variation in cell surface antigens is thus a survival advantage for the bacteria and an essential feature to interact with host cells (11).

In *H. pylori* genomes, there exist two homologous α 1,3/4-FucT genes, *futA* and *futB*, and one gene *futC* for α 1,2-FucT (12). These three *fut* genes do not always encode functional proteins. For instance, *futA*, but not *futB*, encodes an active α 1,3/4-FucT in *H. pylori* strains NCTC11639 and UA948 (13, 14). The *futC* gene in strain NCTC11639 is functional (12) but not functional in strain UA948 (14). The on/off status of *fut* genes and the various levels of FucT activities present in different *H. pylori* strains determine the Lewis antigen expression patterns of *H. pylori* LPS. *FutA* and *FutB* contain a variable number of DD(or N)LRV(or I)NY tandem repeats in the C terminus.

FucTs are inverting glycosyltransferases (GTs), *i.e.* the enzymes catalyze the fucosyl transfer from the donor GDP- β -L-fucose to form an α -glycosidic linkage. On the basis of the CAZy classification, FucTs are categorized into family GT10. Furthermore, FucTs also fall into three sub-families, namely α 1,2-, α 1,3/4-, and α 1,6-FucTs, according to the position of the glycosidic bond each enzyme forms. *H. pylori* α 1,3/4-FucTs are

⁴ The abbreviations used are: LPS, lipopolysaccharide; FucT, fucosyltransferase; GT, glycosyltransferase; LacNAC, *N*-acetyllactosamine; GlcNAC, *N*-acetylglucosamine; AGT, α -glucosyltransferase; BGT, β -glucosyltransferase; MurG, *E. coli* *N*-acetylglucosaminyltransferase; GtfD, TDP-vancosaminyltransferase; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; r.m.s.d., root mean square deviation.



composed of an N-terminal catalytic domain, 2–10 heptad repeats, and a C-terminal tail rich in both basic and hydrophobic residues. The heptad repeats, previously proposed to form a leucine zipper, are essential for homodimerization (15). The positively charged yet hydrophobic tail is required for association with the cell membrane. In contrast, mammalian $\alpha 1,3/4$ -FucTs have the typical structure of type II transmembrane proteins, consisting of a short N-terminal cytoplasmic tail, a transmembrane domain, and a stem region followed by a large C-terminal catalytic domain. The primary sequences of *H. pylori* and mammalian $\alpha 1,3/4$ -FucTs have homology mainly on the nucleotide binding site. For instance, comparison of the *H. pylori* FucT (NCTC11639) polypeptide sequence with mammalian FucTs, including human FucT III to VII, bovine FucTIII, and mouse FucT IV, reveals sequence similarity (40–45% identity) within a stretch of 69 amino acid residues (corresponding to Leu-229 to Leu-297 in *H. pylori* $\alpha 1,3$ -FucT). Therefore, the overall sequence difference makes the *H. pylori* enzyme an attractive target for therapeutic intervention. Despite the biological importance of human and *H. pylori* FucTs, further understanding is impeded by the lack of an x-ray structure.

Recently we and others demonstrated that systematic deletion of the C terminus of *H. pylori* $\alpha 1,3$ -FucT greatly improved the marginal solubility of the full-length protein (15, 16). Up to 80 residues, including the tail containing hydrophobic and positively charged residues (sequence 434–478) and five of the ten heptad repeats (sequence 399–433), can be removed without significant change in structure and catalysis. Several biophysical studies indicate that half of the heptad repeats are essential for the secondary and native quaternary structures (15). We herein present for the first time x-ray crystallographic structures of FucT. The results may suggest a reaction mechanism of FucT and provide a basis for both LPS variation and the design of inhibitors.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification of the Truncated, Selenomethionine-labeled, and Mutant FucTs—The $\alpha 1,3$ -FucT from *H. pylori* strain NCTC11639 was prepared according to the published procedure (15). Protein crystallization was successful with deletion of C-terminal 115 residues. Approximately 160 mg was obtained for crystallization studies, with a purity of >95% as determined by SDS-PAGE. The C-terminal truncated enzyme has only four methionine residues, and their full substitutions with selenomethionine were not sufficient for phase angle calculations. We thus mutated Leu-202, which is located in the middle of a long α -helix based on the secondary structure prediction, to methionine to increase the anomalous difference. The selenomethionine-labeled L202M FucT was produced in *Escherichia coli* by using a non-auxotrophic protocol and purified in a similar manner. The molecular weight

was confirmed by electrospray ionization-mass spectrometry. The enzyme activity was measured based on the formation of GDP that was coupled with the pyruvate kinase/lactate dehydrogenase assay to monitor the consumption of NADH with a fluorescence excitation at 340 nm and emission at 460 nm (15).

Crystallization, Data Collection, and Structure Determination—Purified enzyme was concentrated to 18–20 mg/ml and crystallized at room temperature by the hanging drop vapor diffusion method. Orthorhombic crystals for FucT, selenomethionine-labeled, and mutant enzymes were grown by using equal volumes of the protein solution and the reservoir that contained 4–5% polyethylene glycol 3350, 0.2–0.4 M ammonium sulfate, and 100 mM Bis-Tris (pH 5.5). The substrate-bound FucT complexes were obtained by soaking the apo crystals prior to data collection with a 2–10 mM solution in the crystallization buffer as detailed in Table 1. Before data collection, all crystals were briefly soaked in a crystallization buffer containing 25–30% (v/v) ethylene glycol and were flash frozen in a gaseous stream of liquid nitrogen. The diffraction data were processed and scaled by using HKL/HKL2000 packages (17). Statistics are shown in Table 1. The crystal belongs to the orthorhombic space group $P2_12_12$, with unit cell dimensions $a = 104 \text{ \AA}$, $b = 136 \text{ \AA}$, and $c = 96 \text{ \AA}$, in which an asymmetric unit comprises three FucT molecules.

Using the multiple anomalous dispersion data sets of L202M at 2.5- \AA resolution, the selenium atoms were located, and initial phase angles were calculated by using the program SOLVE (18). Different trials with various parameters in RESOLVE (19, 20) allowed automatic tracings by computer for 60–70% of the amino acid sequence. Statistics for the multiple anomalous dispersion phasing are shown in supplemental Table S1. Because there are three FucT molecules in an asymmetric unit, the completeness of the protein model was improved by cross-referencing of the auto-built polypeptide fragments. The electron densities were also sufficiently clear to allow manual fitting of most missing residues. Using the program O (21), a protomer containing residues 1–348 was thus constructed and placed in all three positions. The model yielded an initial R -value of 0.45 at 2.5- \AA resolution. Further refinement employed the native data set and the program CNS (22), in which 5% reflections were set aside for R_{free} calculation (23). With strong non-crystallographic symmetry restraints, the model gave R and R_{free} of 0.30 and 0.32 at 1.9- \AA resolution.

The electron density map revealed some regions with significant deviations among the three monomers (A, B, and C), which were then excluded from the non-crystallographic symmetry restraints in subsequent refinement. Helix $\alpha 4$ did not have strong densities in monomer A and was not seen in monomer B. Water molecules and sulfate ions were included at later stages. The FucT crystals in complex with GDP and GDP-fu-

FIGURE 1. Overall structure of FucT-GDP-fucose complex. *A*, the complex (stereo view) is shown as a ribbon diagram with the β -strands colored in cyan and yellow for the N- and C-terminal domains, respectively. GDP-fucose is drawn as a ball-and-stick model. *B*, a stereo view of GDP-fucose binding site. The $2F_o - F_c$ electron density maps are contoured at 1σ for the substrate. *C*, the ligands of sulfate (red), GDP (green), and GDP-fucose (blue) bound to the active site are superimposed and presented as stick models in stereo view. Note the outward rotation of the α -phosphate and the shift of the β -phosphate in GDP as compared with GDP-fucose. *D*, partial sequence alignment of *H. pylori* FucT (*hpFucT*) and human FucT (*hFucT*). Identical residues are shaded in gray, and the conserved residues involved in GDP-fucose binding are red. The corresponding secondary structural elements are shown above the sequence.

TABLE 1

Data collection and refinement statistics for the FucT crystals

Numbers in parentheses are for the highest resolution shells. All positive reflections were used in the refinement.

Soaking condition	Native	GDP	GDP-fucose
		2 mM for 4 h	10 mM for 30 min
Data collection			
Beam line	BL12B2, SPring-8	BL12B2, SPring-8	BL13B1, NSRRC
Space group	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2
Unit cell <i>a</i> , <i>b</i> , <i>c</i> (Å)	104.6, 136.2, 96.8	104.2, 136.3, 96.4	104.0, 135.7, 96.2
Resolution (Å)	30-1.9 (1.97-1.90)	30-1.9 (1.97-1.90)	30-2.05 (2.12-2.05)
Number of observations	698,412 (59,587)	618,843 (51,083)	510,822 (38,744)
Unique reflections	107,364 (10,125)	104,965 (9,827)	83,859 (7,798)
Completeness (%)	98.6 (94.1)	96.5 (91.4)	97.9 (92.2)
Average I/σ(I)	35.7 (2.5)	42.7 (3.4)	22.5 (2.1)
R _{merge} (%)	5.8 (86.9)	4.7 (60.3)	8.0 (65.6)
Refinement			
Number of reflections (F > 0)	100,376 (7,963)	96,755 (8,110)	79,019 (6,436)
R _{work} (95% data)	0.191 (0.318)	0.190 (0.283)	0.190 (0.294)
R _{free} (5% data)	0.242 (0.373)	0.237 (0.335)	0.246 (0.329)
r.m.s.d. bond distance (Å)	0.020	0.020	0.020
r.m.s.d. bond angle (°)	1.9	1.9	1.9
Ramachandran plot (% residues)			
In most favored regions	87.5	89.5	88.3
In additional allowed regions	12.2	10.2	11.4
In generously allowed regions	0.3	0.3	0.3
Average B (Å ²)/no. of non-H atoms			
Protein	37.1/8,515	36.1/8,442	40.0/8,480
Water	52.8/1,752	56.9/2,237	53.4/1,306
SO ₄	62.8/15	69.2/5	78.4/10
GDP or GDP-fucose		39.3/84	42.5/114

cose were isomorphous to the apo crystal. Direct use of a previous mid-stage model in rigid-body refinement yielded *R* values of 0.274 and 0.276, respectively, for the FucT·GDP and FucT·GDP-fucose crystals. The initial difference Fourier maps clearly indicated that the pyrophosphate group of GDP replaced the bound sulfate ion in the active site. The sugar moiety of the bound GDP-fucose was visible in all three monomers, although the densities were weaker for monomer A. The FucT·GDP-fucose crystal also allowed slight extension of the C termini, but the amino acids beyond 352 were still invisible.

Synthesis of Enzyme Inhibitors—Preparation of compounds **1** and **2** (see Fig. 5C) followed the procedure reported by Wong and his coworkers (24). Compounds **1** and **2** were individually coupled with 80 different acids (supplemental Fig. S1) to generate the amide derivatives in microplates in the presence of (1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (1 eq) and diisopropyl ethylamine (2 eq) in dimethylformamide. Without purification, the mixtures were subjected to the enzyme activity assay. Compound **3** was synthesized by 1,3-dipolar cyclization (known as “click chemistry”) of propargyl-GDP (the alkyne) and 5-azidopentanoic acid (biphenyl-4-ylmethyl)amide (the azide) in the presence of CuSO₄ and copper wire at room temperature. The solvent mixture was composed of H₂O, EtOH, and t-BuOH (3:2:5) (25). After purification, the resulting product produced ¹H and ¹³C NMR spectra that are consistent with the reported data. These prepared compounds were studied for inhibition in accordance with the aforementioned procedure (15).

RESULTS AND DISCUSSION

Structure of FucT and Its Complexes with GDP/GDP-Fucose—FucT crystallization was only successful upon deletion of the C-terminal 115 residues; this truncated form retained 20% of the native enzyme activity (15). The refined FucT structures all

TABLE 2

H-bonds between GDP/GDP-fucose and FucT residues

FucT residues	H-bond distance of	
	GDP	GDP-fucose
	Å	
Val-222-O	Base N1 (2.6)	Base N1 (2.7)
Lys-223-O	Base N2 (3.0)	Base N2 (3.2)
Ser-188-O _γ	Base N7 (2.8)	Base N7 (2.8)
Val-186-O	Ribose O2' (2.8)	Ribose O2' (2.8)
Lys-225-Nζ	Ribose O3' (3.1)	Ribose O2' (2.7)
Lys-225-Nζ		Ribose O3' (2.9)
Glu-249-Oε2	Ribose O3' (2.8)	Ribose O3' (2.9)
Ser-188-N	α-Phosphate O2A (2.7)	α-Phosphate O1A (2.9)
Asn-189-N	α-Phosphate O2A (2.8)	α-Phosphate O1A (3.0)
Arg-195-NH1	β-Phosphate O3B (2.7)	α-Phosphate O2A (3.2)
Asn-189-Nδ2		β-Phosphate O1B (2.7)
Lys-250-Nζ	β-Phosphate O2B (2.6)	β-Phosphate O2B (3.1)
Arg-195-NH2	β-Phosphate O3B (2.7)	β-Phosphate O3B (2.6)
Asn-240-Nδ2	β-Phosphate O1B (2.6)	Fucose O2 (2.4)
Tyr-246-OH		Fucose O2 (2.8)
Gly-94-O		Fucose O3 (2.5)
Tyr-246-OH		Fucose O3 (2.5)
Glu-249-Oε1		Fucose O4 (2.7)

contain three protein monomers (denoted A, B, and C) in the asymmetric unit cell, starting at Met-1 and terminating at residues 348–352. Each monomer is composed of two similar domains, both having parallel α/β topology of the Rossmann folds (Fig. 1A). The N- and the C-terminal domains encompass residues 20–150 and 160–320, respectively. The first helix α1 (residues 2–13) interacts with the C-terminal domain, whereas the last helix α12 (residues 328–340) interacts with the N-terminal domain. Refinement statistics for the crystals of FucT and two complexes (one with GDP and the other with GDP-fucose) are shown in Table 1.

In the apo FucT crystal, there are three bound sulfate ions, two of which are at identical sites in monomers B and C, and adjacent to the side chain of Arg-195, an essential residue for enzymatic activity (see below). The third sulfate is bound to

TABLE 3

Deviations between FucT and other GT-B models

BGT stands for phage β -glucosyltransferase, MurG denotes *E. coli* N-acetylglucosaminyltransferase, and GtfD represents TDP-vancosaminyltransferase.

Model (PDB code)	N-terminal domain			C-terminal domain			Domain rotation angle
	No. Cas	Identical residues	r.m.s.d.	No. Cas	Identical residues	r.m.s.d.	
BGT (1NZD)	52	10	1.8	130	17	2.0	36
MurG (1NLM)	50	5	1.9	82	4	2.1	29
GtfD (1RRV)	47	2	2.1	86	5	1.9	38

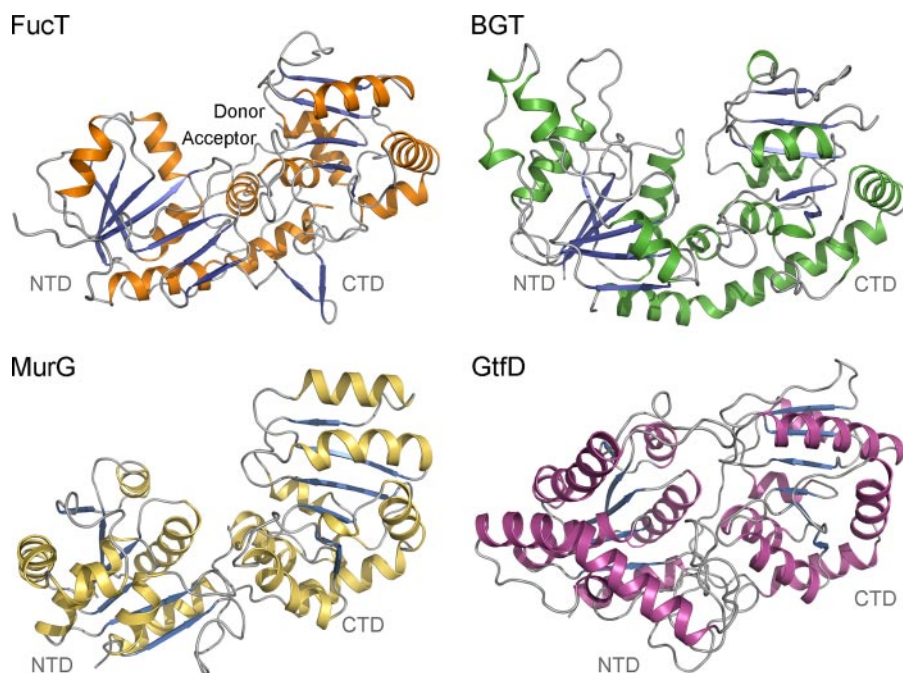


FIGURE 2. Domain arrangements of FucT, BGT, MurG, and GtfD as ribbon diagrams. The C-terminal domains (CTD, right) are oriented according to structural homology, whereas the disposition of individual N-terminal domain (NTD, left) varies.

a distal loop of monomer C. Despite sulfate ions existing in the mother liquor, when the apo crystals were soaked with GDP, the sulfate was readily replaced by GDP. The crystalline FucT retained enzymatic activity to hydrolyze the donor substrate, GDP-fucose in the absence of acceptor. Soaking the FucT crystal with GDP-fucose left only the product GDP in the active site, unless the soaking time was kept to a minimum.

Interactions between the enzyme and the diphosphate are more precisely defined in the GDP-fucose complex of FucT (Fig. 1B) than those for the enzyme and sulfate ion. Both side chains of Arg-195 and Lys-250 provide the neutralizing charges. Among the 18 well defined direct hydrogen bonds (H-bonds) with GDP-fucose (Table 2), there are seven H-bonds between FucT and the nucleoside moiety of GDP, whereas the fucosyl group forms five specific H-bonds with the enzyme. Thus, FucT recognizes its substrate GDP-fucose with very high specificity. GDP alone is bound to the enzyme with a similar repertoire involving 12 direct H-bonds (Table 2), and the diphosphate group is rotated outward (Fig. 1C). In the FucT·GDP-fucose complex structure, one of the two H-bonds of Arg-195 is formed with the α -phosphate and another with the β -phosphate (Fig. 1B). In contrast, both H-bonds from Arg-195 are redirected to the β -phosphate in the FucT·GDP com-

plex. Thus, after the fucose transfer, an additional positive charge is arranged to neutralize the leaving group.

Comparison with Other FucTs—The GDP-fucose binding sites of *H. pylori* and human α 1,3-FucTs have comparatively higher sequence similarity than other regions (Fig. 1D). This region, including the invariant residues Asn-240, Tyr-246, Glu-249, and Lys-250, provides a molecular basis accounting for the same donor-substrate specificity. These residues were studied by site-directed mutagenesis and kinetic analysis to assess their roles in enzyme function (see below).

DALI searches (25) for structural neighbors of FucT provided many Rossmann-fold proteins, including a number of different glycosyltransferases. The conserved N-terminal α - β - α motif in most glycosyltransferases is observed in both domains

of FucT but with little sequence similarity. FucT belongs to the glycosyltransferases-B family in which each protein contains N- and C-terminal domains binding to the acceptor and donor substrates, respectively. To date, there are nine GT families with glycosyltransferases-B folds whose structures are reported (26), including the phage α - and β -glucosyltransferases AGT (family GT 72) (27) and BGT (GT 63) (28), *E. coli* N-acetylglucosaminyltransferase MurG (GT 28) (29), and the vancomycin-synthesizing enzymes GtfA (30), GtfB (31), and GtfD (32) (the three all belonging to GT 1), flavonoid glucosyltransferase (GT 1) (33), glycogen synthase 1 (GT 5) (34), glycosyltransferase AVIGT4 (GT 4), mannosyltransferase wbaZ-1 (GT 4), heptosyltransferase II (GT 9), trehalose-6-phosphate synthase (GT 20) (35, 36), α -2,3-sialyltransferase (GT 80) (37). Among these GT families, GTs 1, 9, 28, 63, and 80 are inverting enzymes like FucTs. The N- and C-terminal domains of FucT can be better superimposed with BGT (PDB code: 1NZD), MurG (PDB code: 1NLM), and GtfD (also called TDP-vancosaminyltransferase; PDB code: 1RRV), as shown in Table 3 and Fig. 2. The central β -sheets match quite well in both domains, whereas the α -helices and peripheral loops show more variations. Each of the two domains has to be superimposed individually because of the different dispositions by 29–38° rotation. The rotations are not

Structure and Mechanism of Fucosyltransferase

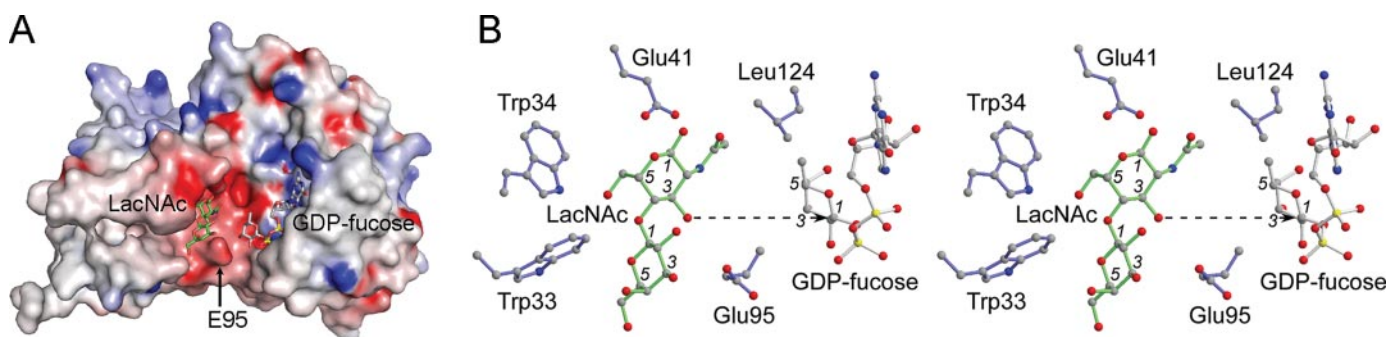


FIGURE 3. Environment of the LacNac model in FucT. *A*, substrate binding of FucT. The molecular surface of FucT is colored from red to blue according to the electrostatic potential from $-15 k_B T$ to $+15 k_B T$ using the program GRASP, where k_B is the Boltzmann constant, and T is the Kelvin temperature. The observed GDP-fucose and predicted LacNac are shown as sticks, in which the carbon bonds are colored in gray and green, respectively. The N- and C-terminal domains correspond roughly to the left and right halves of the surface diagram, respectively. *B*, GDP-fucose, LacNac, and FucT residues are represented with a ball-and-stick model colored gray, green, and blue, respectively (in stereo view). Trp-33, Trp-34, and Glu-41 belong to the $\alpha 1$ - $\beta 1$ - $\alpha 2$ motif. The dashed arrow represents the nucleophilic attack from the C3-OH of GlcNAc to the anomeric position of fucose.

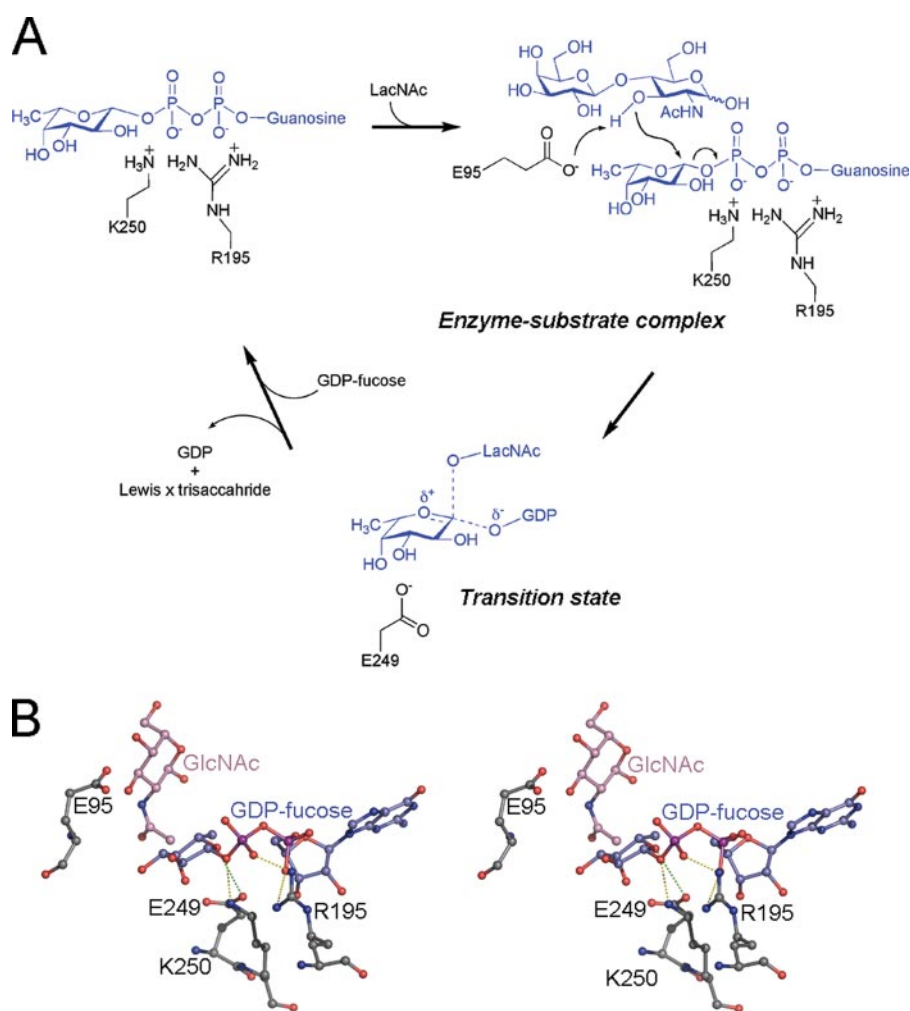


FIGURE 4. FucT active site and proposed reaction mechanism. *A*, proposed catalytic mechanism of FucT where Glu-95 serves as a general base. The developing positive charge on fucose is stabilized in part by Glu-249. *B*, a close-up stereo view of the FucT catalytic region showing the predicted LacNac model (pink), GDP-fucose (blue), and the related FucT residues (gray), which are drawn as stick models. For clarity, only GlcNAc is displayed. Hydrogen bonds are represented as yellow dotted lines.

about a uniform axis, yet the overall conformation of FucT is apparently more open than all other models.

In general, the C-terminal domains are more conserved than the N-terminal domains (Table 3 and Fig. 2). In particular, the C-terminal domains of FucT and BGT have 130 matched pairs

of $C\alpha$ atoms, including 17 identical residues. There are also 10 identities in the N-terminal domain. These correspond to <10% of the entire sequence. The other two models of MurG and GtfD have more variations.

Catalytic Mechanism—BGT, being a member of the glycosyltransferases-B family, uses Asp-100 as the general base. Interestingly, this residue is equivalent to Glu-95 of FucT. Glu-95 is located in the N-terminal domain that presumably associates with the acceptor substrate. The side chain of Glu-95 is positioned immediately adjacent to the anomeric carbon of fucose in the GDP-fucose complex (Fig. 3), which supports its suggested role as a general base in catalysis.

A model of the acceptor *N*-acetylglucosamine (LacNAc) can be constructed based on the locations of the catalytic base and the donor substrate, as well as the shape and orientation of the active site cleft. The shape of LacNAc is complementary to the cleft (Fig. 3A). The 3-OH of *N*-acetylglucosamine (GlcNAc) is close to the anomeric carbon of the fucose residue, as well as the carboxylate of Glu-95 (Fig. 3B). The methyl group of the *N*-acetyl group, *trans* to the C2 of GlcNAc, also makes hydrophobic contact with the side chain of Leu-124. On the distal side, other OH groups may form additional H-bonds with the side chains of the first α - β - α motif of the N-terminal domain, such as Trp-33, Trp-34, and Glu-41 (Fig. 3B).

The structures of the substrate- and product-bound forms of FucT allow us to suggest a catalytic mechanism that is similar to

that of human FucT (38). LacNAc is bound to the pocket in the N-terminal domain, which has a highly negative electrostatic potential (Fig. 3A) that reduces the pK_a of the C3-OH group of GlcNAc to favor nucleophilic attack. Upon deprotonation of the C3-OH group by Glu-95 (Fig. 4A), the acceptor nucleophile can attack the anomeric position of GDP-fucose to form a new glycosidic bond with an inverted configuration (Fig. 4A). The side product GDP dissociates at the same time. Because the anomeric car-

bon is located between the acceptor nucleophile, the C3-OH of GlcNAc, and the leaving group GDP (Fig. 4B), the geometry is consistent with the in-line displacement mechanism (39).

Meanwhile, it is intriguing that two neighboring regions of opposite charges are also observed. The positively charged pocket, in close association with GDP-fucose, facilitates glycosidic bond cleavage by neutralizing the negative charges on GDP. Lys-250 and Arg-195 are the two critical residues (Fig.

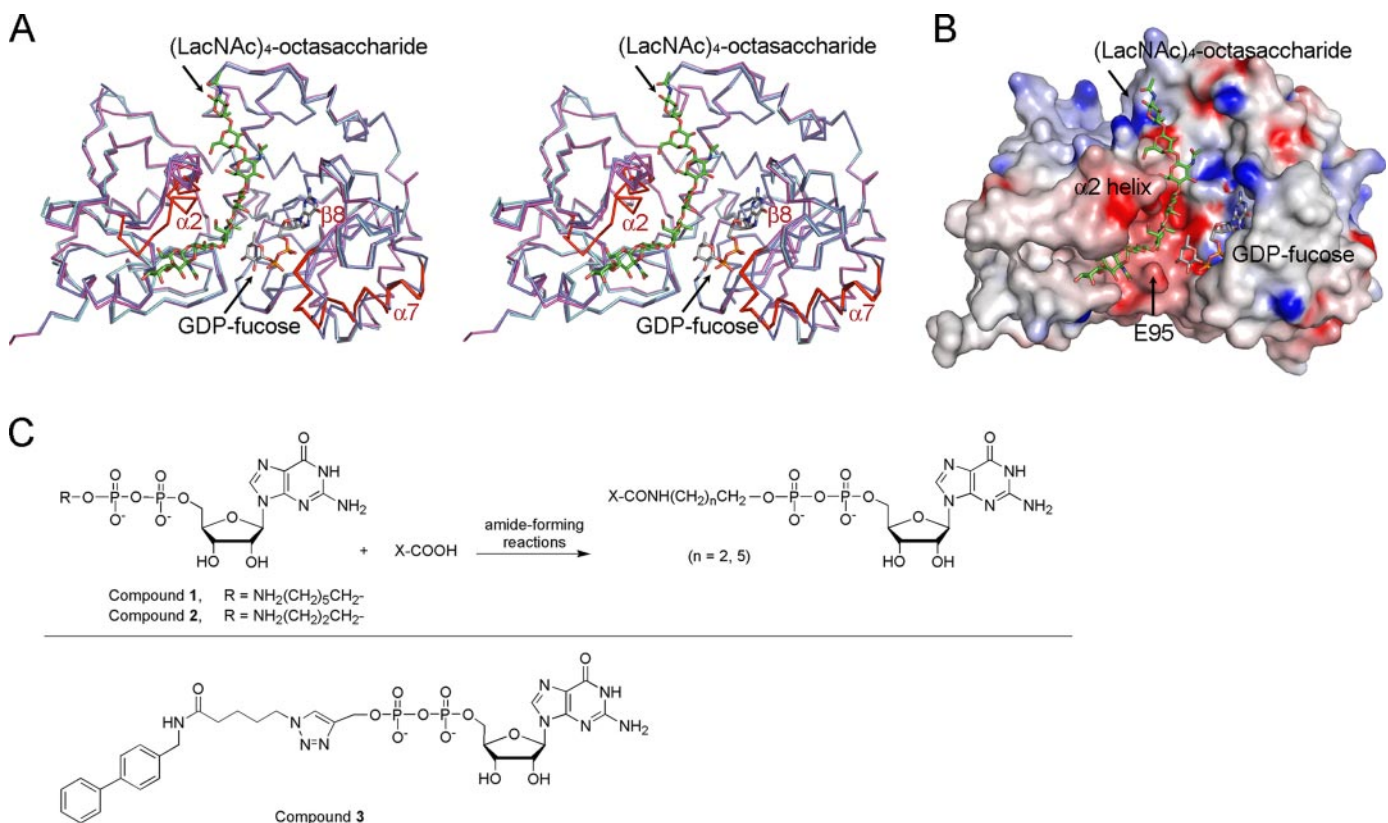


FIGURE 5. Acceptor substrate model, proposed reaction mechanism, and potential FucT inhibitors. *A*, the $C\alpha$ tracings of the FucT monomers in different crystals (in stereo view). Monomers *A*, *B*, and *C* are colored *blue*, *cyan*, and *magenta*, respectively. The bound GDP-fucose is shown as a *stick model* with the carbon backbone colored in *gray*. Two regions of $\alpha 2$ - $\beta 2$ and $\beta 8$ - $\alpha 7$, which undergo conformation changes, are colored *red*. *B*, the molecular surface of the FucT is colored from *red* to *blue* according to the electrostatic potential from $-15 k_B T$ (acidic) to $+15 k_B T$ (basic) using the program GRASP. The observed GDP-fucose and the modeled (LacNAc)₄-octasaccharide ($\beta 1,3$ -linkage between each LacNAc unit) are shown in *gray* and *green*, respectively, for the carbon-carbon bonds. The octasaccharide lies in the crevice between the two binding domains and makes primary contact with the $\alpha 2$ helix. *C*, structures of the potential FucT inhibitors. Compounds **1** and **2** were coupled with 80 different carboxylic acids (X-COOH) by amide-forming reactions to give GDP derivatives that were unpurified and directly subjected to the FucT activity assay. Compound **3** was found to be an inhibitor against *H. pylori* FucT ($K_i = 0.59 \mu M$).

TABLE 4

Kinetic analysis of FucTs and mutants

FucT Δ C45 stands for the FucT to be truncated C-terminal 45 residues, and FucT Δ C115 represents FucT to be truncated C-terminal 115 residues.

FucT or mutant	Relative activity	GDP-fucose			LacNAc		
		K_m	V_{max}	V_{max}/K_m	K_m	V_{max}	V_{max}/K_m
	%	μM	$\mu mol/min/mg$		mM	$\mu mol/min/mg$	
FucT Δ C45	100	88.31 \pm 8.68	5.32 \pm 0.14	0.060	0.71 \pm 0.06	5.28 \pm 0.14	7.44
FucT Δ C115	6.6	222.3 \pm 32.2	1.35 \pm 0.06	0.006	8.92 \pm 1.61	1.48 \pm 0.08	0.17
E95A-FucT Δ C45	ND ^a	—	—	—	—	—	—
E95D-FucT Δ C45	ND	—	—	—	—	—	—
R195A-FucT Δ C45	ND	—	—	—	—	—	—
R195K-FucT Δ C45	3.6	591.2 \pm 41.4	0.96 \pm 0.019	0.002	10.35 \pm 0.67	1.10 \pm 0.027	0.11
N240A-FucT Δ C45	10	80.16 \pm 10.23	1.09 \pm 0.024	0.014	12.77 \pm 0.73	1.78 \pm 0.033	0.14
Y246A-FucT Δ C45	20	>800	—	—	2.39 \pm 0.14	0.28 \pm 0.015	0.12
Y246F-FucT Δ C45	4	157.1 \pm 18.7	0.15 \pm 0.008	0.001	1.29 \pm 0.113	0.17 \pm 0.009	0.13
E249A-FucT Δ C45	ND	—	—	—	—	—	—
E249D-FucT Δ C45	ND	—	—	—	—	—	—
E249Q-FucT Δ C45	ND	—	—	—	—	—	—
K250A-FucT Δ C45	ND	—	—	—	—	—	—

^a ND, not detected because of no or very low level activity (<0.5%). Thus, the kinetic parameters could not be obtained.

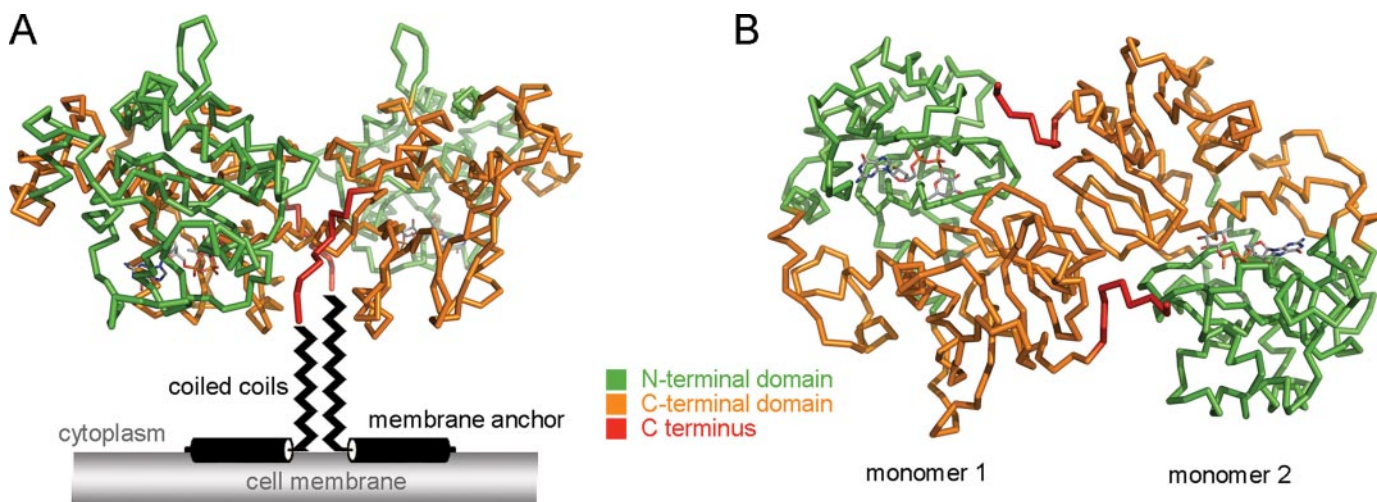


FIGURE 6. **The putative FucT dimer.** *A*, in the dimer, the two C termini have the same orientation. This is consistent with the parallel leucine-zipper coiled-coil as predicted for the full-length enzyme. *B*, the enzyme is turned 90°, viewed along the dyad axis and shown. The domains in each monomer are color-coded as green (N-terminal domain), orange (C-terminal domain), and red (C terminus). GDP-fucose is shown as a stick model.

4*B*). A major conformational change of the Arg-195-containing helix $\alpha 7$ is induced by the binding of GDP or GDP-fucose (Fig. 5*A*). The negatively charged surface flanking the fucose moiety stabilizes the developing oxonium ion in the transition state. In particular, Glu-249 is the key residue near C1 and O5 of fucose at a distance of 4.9 and 3.9 Å (Fig. 4*B*), respectively. On the other hand, a considerable distance (8 Å) exists between the donor and acceptor. Thus, the two domains must move, triggered by the binding of both substrates, to yield a more closed conformation of the protein and facilitate additional cross-domain interactions.

Essential Catalytic Residues Supported by Site-directed Mutagenesis—The binding of donor substrate in the absence of acceptor suggests that FucT catalyzes the sugar transfer reaction by a sequential mechanism, as do most other GTs (26). Site-directed mutants of FucT were thus prepared to support the roles of the key residues identified above. The enzyme activity was measured based on the formation of the side product GDP (15). Although the enzyme is able to hydrolyze GDP-fucose as aforementioned, our isotope-based TLC analysis indicated that the rate of hydrolysis is at least 340-fold lower than that of the enzymatic fucosylation (data not shown). The truncated FucT Δ C115 (representing the protein short of C-terminal 115 residues) has much less activity in comparison with the larger FucT Δ C45 (short of C-terminal 45 residues) that contains an additional leucine zipper in the C terminus and forms a dimer in solution as the full-length FucT (15). All the mutations are based on FucT Δ C45, as shown in Table 4. Substitution of Glu-95 with alanine (E95A) or aspartic acid (E95D) led to complete activity loss, in agreement with the suggested role as a general base. Alanine mutants of Arg-195 or Lys-250 resulted in no detectable activity, supporting the idea that the two residues provide positive charges to interact with negatively charged GDP-fucose (Figs. 1*B* and 4*B*). Glu-249 functions to stabilize the proposed oxonium cation, as well as to form two H-bonds with both the ribose and the fucose residues of GDP-fucose (Figs. 1*B* and 4*B*). The importance of Glu-249 is further strengthened by undetectable activities of the mutants E249A,

E249D, and E249Q. Tyr-246 is involved in the binding with GDP-fucose (Fig. 1*B*), as suggested by the reduced activity of Y246A and the greatly increased K_m GDP-fucose value ($>800 \mu\text{M}$). The mutant N240A retained 10% activity of the wild-type enzyme. Surprisingly, the K_m LacNAc of N240A was 12.8 mM, 18-fold higher than that of the wild-type FucT, whereas the K_m GDP-fucose was not affected. Despite the sequence identity of Asn-240 among various $\alpha 1,3/4$ -FucTs (data not shown), its H-bond to fucose is possibly (Fig. 1*B*) not crucial for the donor substrate binding. By contrast, Asn-240 is implicated in binding with the acceptor substrate. The interaction likely occurs either after the conformation change induced by the occupancy of both donor and acceptor substrates, or in the larger proteins (with less C-terminal truncation).

LPS Model and Inhibitor Design—Because FucT catalyzes the fucosyl transfer to oligosaccharyl/polysaccharyl LacNAc-containing LPS *in vivo*, a model of the (LacNAc) $_4$ -octasaccharide ($\beta 1,3$ -linkage between each LacNAc unit) was constructed (Fig. 5*B*) based on the aforementioned considerations. The modeled structure indicates that the octasaccharide lies in the crevice between the two binding domains, in proximity to the $\alpha 2$ helix (Fig. 5). By coincidence, helix $\alpha 2$ adopts different conformations among the three structures (Fig. 5*A*), which may provide better contact with the LacNAc acceptor. Meanwhile, at least 4–5 LacNAc units are encompassed by the enzyme (Fig. 5), in agreement with the need to accommodate a polysaccharide acceptor for fucosylation.

Because FucT has more interactions with GDP than it does with fucose, it would be necessary to mimic the former interactions when developing FucT inhibitors. A variety of molecules were prepared by coupling of GDP-hexylamine (compound 1, Fig. 5*C*) with 80 different carboxylic acids (supplemental Fig. S1). The inhibitory activity of the resulting unpurified products was assessed using the FucT enzyme assay. Most of the molecules had IC_{50} values of 10–100 μM , comparable to that of GDP. Similar results were obtained using the molecules derived from the reactions of GDP-propylamine (compound 2) and carboxylic acids. Compound 3, reported previously as a potent inhib-

itor of human FucT VI (24), effectively inhibited *H. pylori* FucT ($K_i = 0.59 \mu\text{M}$). This lead compound represents a 150-fold enhanced affinity compared with GDP-fucose ($K_m = 88 \mu\text{M}$). Because there is considerable distance between the donor and acceptor sites, the improved inhibition is likely due to the substantial molecular size that makes it possible to gain extra multiple interactions with the acceptor binding site.

Possible Ramifications of an Intertwining Dimer—One asymmetric unit of the FucT crystal contains three molecules of the protein. Monomers A and B are related by a non-crystallographic 2-fold axis and constitute a possible dimer. A similar dimer is formed by monomer C and the adjacent monomer C' related by a crystallographic dyad. Additionally the remaining C termini of both monomers are oriented in the same direction (Fig. 6A) and would probably be followed by the leucine zipper motifs (truncated in our experimental protein) in the full-length protein. As a result, the dimer observed is consistent with the proposed model (15) in which the C-terminal heptads of each subunit interact with those of the other subunit in parallel to yield a dimer. Although the studied enzyme lacks the C-terminal heptad repeats plus a membrane anchor and appears to be a monomer in solution (15), the structural information suggests that the homodimer in our crystal is not entirely caused by crystal packing. Rather, it may present a native form of the enzyme.

In the dimeric FucT structure, the C terminus of the counter subunit is 15–18 Å away from the predicted catalytic base Glu-95, whereas the residue is much farther away (32–33 Å) from that of the same subunit (Fig. 6B). Owing to the important role of the C terminus in the acceptor binding and specificity (40), the C terminus (residues 350–363) that is missing in the resolved x-ray structures likely crosses over the interface to approach the active site of the counter monomer and thus participates in acceptor binding. This hypothesis is supported by the observation that the truncation of the heptad repeats has greater impact on acceptor binding than donor binding (Table 4).

A molecular ruler mechanism of FucT was proposed for how *H. pylori* varies its LPS fucosylation pattern (41). The formation of a homo- and heterodimeric FucT, due to the expression of the genes *futA* and *futB*, is putatively linked to LPS variation. Because the region of heptad repeats mediates the formation of homo- and heterodimers, a number of different dimers, depending on the number and identity of the heptads involved in dimerization, may affect the sizes of O-antigen fucosylation units. This mechanism thus suggests that a homodimer preferentially fucosylates one LacNAc unit. Nevertheless, our preliminary study on the fucosylation of oligomeric LacNAc indicates that the FucT homodimer is able to carry out multiple fucosylations (data not shown), suggesting that the ruler mechanism requires re-evaluation. In accordance with our structures, the enzyme not only accommodates four to five LacNAc units but also exists as an intertwining dimer, the features of which are likely applicable to the formation of heterodimers: only one or two isoforms can be generated due to head-to-head interactions of the two catalytic domains.

In summary, the first x-ray structures of FucT suggest a reaction mechanism at the molecular level. This mechanism is help-

ful in the design of potential lead compounds for enzyme inhibition. Further optimization of inhibitors and crystallization of FucT/inhibitor complexes will potentiate the development of a new treatment for *H. pylori* infection.

Acknowledgments—We thank the National Synchrotron Radiation Research Center of Taiwan, and SPring-8 of Japan for beam time allocation, Hui-Ling Shr at the core facility of X-ray crystallography, the Institute of Biological Chemistry (IBC), Academia Sinica, for crystal screening and refinement, Dr. Yi-Sheng Cheng (Institute of Molecular Biology) and I-Lin Wu (IBC, Academia Sinica) for technical assistance, and the National Core Facility of Proteomics for mass spectrometry analysis.

REFERENCES

- Graham, D. Y. (1991) *J. Gastroenterol. Hepatol.* **6**, 105–113
- Peterson, W. L. (1991) *N. Engl. J. Med.* **324**, 1043–1048
- Parsonnet, J. (1996) *N. Engl. J. Med.* **335**, 278–280
- Nakamura, S., Yao, T., Aoyagi, K., Iida, M., Fujishima, M., and Tsuneyoshi, M. (1997) *Cancer* **79**, 3–11
- Ilver, D., Arnqvist, A., Ogren, J., Frick, I. M., Kersulyte, D., Incecik, E. T., Berg, D. E., Covacci, A., Engstrand, L., and Boren, T. (1998) *Science* **279**, 373–377
- Mahdavi, J., Sonden, B., Hurtig, M., Olfat, F. O., Forsberg, L., Roche, N., Angstrom, J., Larsson, T., Teneberg, S., Karlsson, K. A., Altraja, S., Wadstrom, T., Kersulyte, D., Berg, D. E., Dubois, A., Petersson, C., Magnusson, K. E., Norberg, T., Lindh, F., Lundsberg, B. B., Arnqvist, A., Hammarstrom, L., and Boren, T. (2002) *Science* **297**, 573–578
- Rathbone, B. J., Wyatt, J. I., Worsley, B. W., Shires, S. E., Trejdosiewicz, L. K., Heatley, R. V., and Losowsky, M. S. (1986) *Gut* **27**, 642–647
- Wirth, H. P., Yang, M., Karita, M., and Blaser, M. J. (1996) *Infect. Immun.* **64**, 4598–4605
- Monteiro, M. A., Chan, K. H., Rasko, D. A., Taylor, D. E., Zheng, P. Y., Appelmelk, B. J., Wirth, H. P., Yang, M., Blaser, M. J., Hynes, S. O., Moran, A. P., and Perry, M. B. (1998) *J. Biol. Chem.* **273**, 11533–11543
- Lozniewski, A., Haristoy, X., Rasko, D. A., Hatier, R., Plenat, F., Taylor, D. E., and Angioi-Duprez, K. (2003) *Infect. Immun.* **71**, 2902–2906
- Hallet, B. (2001) *Curr. Opin. Microbiol.* **4**, 570–581
- Wang, G., Rasko, D. A., Sherburne, R., and Taylor, D. E. (1999) *Mol. Microbiol.* **31**, 1265–1274
- Ge, Z., Chan, N. W., Palcic, M. M., and Taylor, D. E. (1997) *J. Biol. Chem.* **272**, 21357–21363
- Rasko, D. A., Wang, G., Palcic, M. M., and Taylor, D. E. (2000) *J. Biol. Chem.* **275**, 4988–4994
- Lin, S.-W., Yuan, T.-M., Li, J.-R., and Lin, C.-H. (2006) *Biochemistry* **45**, 8108–8116
- Ma, B., Audette, G. F., Lin, S., Palcic, M. M., Hazes, B., and Taylor, D. E. (2006) *J. Biol. Chem.* **281**, 6385–6394
- Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307–326
- Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. D. Biol. Crystallogr.* **55**, 849–861
- Terwilliger, T. C. (2000) *Acta Crystallogr. D. Biol. Crystallogr.* **56**, 965–972
- Terwilliger, T. C. (2003) *Acta Crystallogr. D. Biol. Crystallogr.* **59**, 38–44
- Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard (1991) *Acta Crystallogr. A* **47**, 110–119
- Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D. Biol. Crystallogr.* **54**, 905–921
- Brunger, A. T. (1993) *Acta Crystallogr. D. Biol. Crystallogr.* **49**, 24–36
- Lee, L. V., Mitchell, M. L., Huang, S. J., Fokin, V. V., Sharpless, K. B., and Wong, C. H. (2003) *J. Am. Chem. Soc.* **125**, 9588–9589
- Holm, L., and Sander, C. (1993) *J. Mol. Biol.* **233**, 123–138
- Qasba, P. K., Ramakrishnan, B., and Boeggeman, E. (2005) *Trends Biochem. Sci.* **30**, 53–62

Structure and Mechanism of Fucosyltransferase

27. Lariviere, L., Sommer, N., and Morera, S. (2005) *J. Mol. Biol.* **352**, 139–150
28. Vrielink, A., Ruger, W., Driessen, H. P., and Freemont, P. S. (1994) *EMBO J.* **13**, 3413–3422
29. Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., Mokhtarzadeh, M., and Walker, S. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 845–849
30. Mulichak, A. M., Losey, H. C., Lu, W., Wawrzak, Z., Walsh, C. T., and Garavito, R. M. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 9238–9243
31. Mulichak, A. M., Losey, H. C., Walsh, C. T., and Garavito, R. M. (2001) *Structure* **9**, 547–557
32. Mulichak, A. M., Lu, W., Losey, H. C., Walsh, C. T., and Garavito, R. M. (2004) *Biochemistry* **43**, 5170–5180
33. Offen, W., Martinez-Fleites, C., Yang, M., Kiat-Lim, E., Davis, B. G., Tarling, C. A., Ford, C. M., Bowles, D. J., and Davies, G. J. (2006) *EMBO J.* **25**, 1396–1405
34. Buschiazzo, A., Ugalde, J. E., Guerin, M. E., Shepard, W., Ugalde, R. A., and Alzari, P. M. (2004) *EMBO J.* **23**, 3196–3205
35. Gibson, R. P., Tarling, C. A., Roberts, S., Withers, S. G., and Davies, G. J. (2004) *J. Biol. Chem.* **279**, 1950–1955
36. Gibson, R. P., Turkenburg, J. P., Charnock, S. J., Lloyd, R., and Davies, G. J. (2002) *Chem. Biol.* **9**, 1337–1346
37. Ni, L., Sun, M., Yu, H., Chokhawala, H., Chen, X., and Fisher, A. J. (2006) *Biochemistry* **45**, 2139–2148
38. Murray, B. W., Wittmann, V., Burkart, M. D., Hung, S. C., and Wong, C. H. (1997) *Biochemistry* **36**, 823–831
39. Pedersen, L. C., Darden, T. A., and Negishi, M. (2002) *J. Biol. Chem.* **277**, 21869–21873
40. Ma, B., Wang, G., Palcic, M. M., Hazes, B., and Taylor, D. E. (2003) *J. Biol. Chem.* **278**, 21893–21900
41. Nilsson, C., Skoglund, A., Moran, A. P., Annuk, H., Engstrand, L., and Normark, S. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103**, 2863–2868