Structural Basis of the Catalytic Reaction Mechanism of Novel 1,2-α-L-Fucosidase from Bifidobacterium bifidum

Masamichi Nagae, Atsuko Tsuchiya, Takane Katayama, Kenji Yamamoto, Soichi Wakaatsuki, and Ryuichi Kato

From the Structural Biology Research Center, Photon Factory, Institute of Materials Structure Science, High Energy Accelerator Research Organization (KEK), Tsukuba, Ibaraki 305-0801, the Division of Integrated Life Science, Graduate School of Biostudies, Kyoto University, Kitashirakawa, Sakyo-ku, Kyoto 606-8502, and Research Institute for Bioresources and Biotechnology, Ishikawa Prefectural University, Nonoi, Ishikawa 921-8836, Japan

Received for publication, March 15, 2007, and in revised form, April 11, 2007 Published, JBC Papers in Press, April 25, 2007, DOI 10.1074/jbc.M702246200

1,2-α-L-Fucosidase (AfcA), which hydrolyzes the glycosidic linkage of Fucα1–2Gal via an inverting mechanism, was recently isolated from Bifidobacterium bifidum and classified as the first member of the novel glycoside hydrolase family 95. To better understand the molecular mechanism of this enzyme, we determined the x-ray crystal structures of the AfcA catalytic (Fuc) domain in unliganded and complexed forms with deoxy-fuconojirimycin (inhibitor), 2′-fucosyllactose (substrate), and l-fucose and lactose (products) at 1.12–2.10 Å resolution. The AfcA Fuc domain is composed of four regions, an N-terminal region, a helical linker, an (α/α)6 helical barrel domain, and a C-terminal β region, and this arrangement is similar to bacterial phosphorylases. In the complex structures, the ligands were buried in the central cavity of the helical barrel domain. Structural analyses in combination with mutational experiments revealed that the highly conserved Glu566 probably acts as a general acid catalyst. However, no carboxylic acid residue is found at the appropriate position for a general base catalyst. Instead, a water molecule stabilized by Asn423 in the substrate-bound complex is suitably located to perform a nucleophilic attack on the C1 atom of l-fucose moiety in 2′-fucosyllactose, and its location is nearly identical near the O1 atom of β-1-fucose in the products-bound complex. Based on these data, we propose and discuss a novel catalytic reaction mechanism of AfcA.

Bifidobacteria are obligate anaerobic lactic acid-producing bacteria that constitute a substantial fraction of the mammalian intestinal microflora. Interest in bifidobacteria has grown recently following the identification of many beneficial probiotic homeostatic effects, including reducing the presence of harmful bacteria and toxic compounds, immunomodulation, and anticarcinogenic activity (1–5). Bifidobacteria naturally colonize the lower gastrointestinal tract, but host and microbial factors present in the upper gastrointestinal tract render this environment poor in mono- and disaccharides. Therefore, to survive here, bifidobacteria have evolved the ability to produce a variety of different surface-bound or secreted glycosidases facilitating the metabolism of the diverse sugars found in the lower gastrointestinal tract (6–8).

There are two broad classes of glycosidases, exo- and endoenzymes, that differ based on the site and mechanism of polysaccharide degradation. Endoglycosidases cleave specific internal glycosidic bonds, but exoglycosidases remove oligosaccharide units at the reducing or nonreducing ends of the polysaccharide chain. Glycosidases can be further divided into two broad families, retaining and inverting glycosidases, according to the stereochemical outcome of their action (9, 10). Most retaining glycosidases have two catalytic carboxylic acids separated by 5.5 Å in their active site and function through a double displacement mechanism (9–11). In contrast, inverting glycosidases act through a single displacement mechanism in which the two carboxyl groups, acting as general acid and base catalysts, are ~10.5 Å apart, allowing simultaneous interactions with a water molecule and substrate (9–11). Glycosidic cleavage involves the protonation of the glycosidic oxygen by the general acid catalyst in concert with general base-catalyzed nucleophilic attack of a water molecule at the anomeric center. The result is a hemiacetal product with an anomeric configuration that is inverted relative to that of the substrate.

1,2-α-L-Fucosidase (AfcA) is an exoglycosidase recently identified from Bifidobacterium bifidum (12). It was classified as a member of the glycoside hydrolase (GH) family 95 using the CAZy server (available on the World Wide Web at www.cazy.org/) (13). AfcA has 1,959 amino acids divided among three domains: an N-terminal domain with unknown function, a catalytic domain (Fuc domain), and a C-terminal bacterial Ig-like domain. The recombinantly expressed Fuc domain exhibits 1,2-α-L-fucosidase (EC 3.2.1.63) activity leading to the removal of α1–2 fucosyl residues from the nonreducing ends of oligosaccharides, such as 2′-fucosyllactose (2′FL; Fucα1–2Galβ1–4Glc) and lacto-N-fucopentaose I (Fucα1–2Galβ1–3GlcNAcβ1–3Galβ1–4Glc) (12). The stereochemical outcome of the released l-fucose was determined to be inversion by 1H NMR (12).

The abbreviations used are: AfcA, 1,2-α-L-fucosidase; GH, glycoside hydrolase; 2′FL, 2′-fucosyllactose; DFJ, deoxy-fuconojirimycin; MES, 2-N-morpholinooethanesulfonic acid; PEG, polyethylene glycol; r.m.s., root mean square.
Crystal Structure of Afca Catalytic Domain

L-Fucose is a common monosaccharide present at the non-reducing ends of oligosaccharides of many glycoconjugates, including N- and O-linked oligosaccharides of glycoproteins, glycolipids on the cell surface, blood group substances, and oligosaccharides in human milk (14, 15). A number of different species from bacteria to mammals express α-L-fucosidases, and the abundance of these enzymes emphasizes the ubiquity and biological significance of L-fucose in living organisms. Fucose-containing carbohydrates play important roles in diverse cellular and physiologic processes, including inflammatory responses (16–18) and antigenic determination (19). Furthermore, altered expression of α-L-fucosidases and corresponding changes in fucosylation levels are seen in many carcinomas (20, 21). Accordingly, the presence of human specific L-fucose residues.

To date, most identified α-L-fucosidases belong to GH family 29, based on amino acid sequence similarities, and the only solved structure from this family is the Thermotoga maritima α-L-fucosidase (25). However, this is a retaining glycosidase, and the structural characteristics of inverting α-L-fucosidases remain unclear. To better understand the architecture and mechanism of inverting α-L-fucosidases, we performed structural and biochemical studies of the B. bifidum 1,2-α-L-fucosi-
dase catalytic domain. A series of the structural and the bio-
chemical analyses provides insight into a novel catalytic mechanism for this enzyme.

EXPERIMENTAL PROCEDURES

Materials—Q Sepharose Fast Flow, CHT ceramic hydroxyapatite, cellulofine GC-700-m, and MonoQ columns were purchased from Amersham Biosciences, Bio-Rad, Seikagaku Kogyo, and Amersham Biosciences, respectively. All crystallization reagents were purchased from Hampton Research and deCODE Genetics. Deoxyfuconojirimycin (DFJ) was purchased from Seikagaku Kogyo (Japan), and 2’FL, L-fucose, and lactose were from Sigma. Other chemicals were obtained from Wako Pure Chemical (Japan), Sigma, and Nacalai Tesque (Japan).

Protein Expression and Purification—The gene corresponding to the catalytic domain (Fuc domain; 577–1,474 amino acid residues of full-length Afca fucosidase) was amplified by PCR using pSA3 as a template (12) and a pair of primers (forward, 5’-CCATATGGTCATCGCCAGTGTCGAGGACG-3’; reverse, 5’-GCCCGGGTCAGCTCGCCTTCTTCGTGATCG-3’). After digestion with Ndel and Smal, the amplified fragment was inserted into the Ndel-BamHI (blunt-ended) site of pET-3a (Novagen). Escherichia coli strain BL21(DE3) cells carrying the resultant plasmid were grown in LB medium and induced by the addition of 1 mM isopropyl-β-D-thiogalactopyranoside at 291 K for 35 h. Cells were disrupted by sonication at 277 K. The supernatant was fractionated by the addition of 1.25 M sodium azide (dissolved in 100 mM sodium phosphate buffer, pH 7.0) to a reaction mixture consisting of 8.5 microunits of fucose dehydrogenase from a Hogland strain (28). Initial velocity of hydrolysis was determined in the range of 0–200 mM NaCl. Appropriate fractions were precipitated again with 80% saturated ammonium sulfate and dialyzed against 2 mM potassium phosphate buffer, pH 7.0, and 20 mg/liter phenylmethanesulfonyl fluoride. The resulting sample was applied onto a CHT ceramic hydroxyapatite column. After the elution with a linear gradient of 2–200 mM potassium phosphate buffer, pH 7.0, enzymatically active fractions were collected and loaded onto a gel filtration column (cellulofine GC-700-m) equilibrated in 10 mM sodium phosphate (pH 7.0), 100 mM NaCl, 1 mM EDTA, and 20 mg/liter phenylmethanesulfonyl fluoride. The Fuc domain was finally purified on a MonoQ column and concentrated to 10 mg/ml in 10 mM Tris-HCl (pH 8.0). Selenomethionine-substituted protein was produced using E. coli strain B834(DE3) in LeMaster medium containing 25 mg/liter seleno-l-methionine (26, 27). Purification of the selenium-labeled protein was carried out as described above for the native Fuc domain without the gel filtration step. Mass spectrometric analysis (Autoflex; Bruker-Daltonics) verified that all 12 methionine residues in the 898-amino acid Fuc domain were substituted with selenomethionine. The native and selenium-labeled proteins were purified as a single band on SDS-polyacrylamide gel electrophoresis stained with Coomassie Brilliant Blue.

Site-directed Mutagenesis—Amino acid substitutions of the Afca Fuc domain were introduced by site-directed mutagenesis involving PCR. Mutant proteins for crystallographic study were expressed and purified as described above. Those for biochemical analyses were expressed as C-terminal hexahistidine-tagged proteins using modified pET-3a vectors and were purified by the MagneHis purification system (Promega) and a MonoQ column.

Inductively Coupled Plasma Emission Spectroscopy—The purified Fuc domain was dialyzed against 10 mM Tris-HCl (pH 6.5) prior to analysis. The content and concentration of metal ions in the protein were determined by inductively coupled plasma emission spectroscopy using ICPS-8000 (Shimadzu). ICP multielement standard solution IV (Merck) was used as a standard.

Enzymatic Assay and Chemical Rescue—Examination of the enzymatic 1,2-α-L-fucosidase activity of the wild-type and mutated proteins was performed at 303 K in 10 mM sodium phosphate buffer (pH 7.0) using 2’FL as a substrate. After the reaction, reaction mixtures were boiled for 3 min. The amount of released L-fucose was determined by a previously described method using fucose dehydrogenase from a Pseudomonas strain (28). Initial velocity of hydrolysis was determined in the range where the linearity of reaction rate was observed. Kinetic parameters were determined by double-reciprocal plot of Michaelis-Menten equation, in which substrate concentration was varied in the range of 0.4–3 times their K_m values.

Chemical rescue experiments were performed by adding 1.25 mM sodium azide (dissolved in 100 mM sodium phosphate, pH 7.0) to a reaction mixture consisting of 8.5 micromoles of wild-type and mutated proteins, 100 mM sodium phosphate (pH 7.0), and 2 mM 2’FL at 303 K. The reaction mixture was then spotted onto a silica gel plate (Merck). The plate was
Crystal Structure of AfcA Catalytic Domain

The apo form of Fuc domain crystals was grown using the hanging drop vapor diffusion method from drops containing an equal volume of protein (10 mg/ml) in 10 mM Tris-HCl (pH 8.0) and precipitant composed of 0.1 M Tris-HCl (pH 8.0), 10% (v/v) isopropyl alcohol, and 10% (w/v) polyethylene glycol (PEG) 4000 at 289 K over 10 days. The selenomethionine-substituted crystals were obtained under almost identical conditions (0.1 M Tris-HCl (pH 7.0), 10% (v/v) isopropyl alcohol, 10% (w/v) PEG 4000). Crystals of the complex of the Fuc domain with the inhibitor DFJ were obtained by soaking the apo form crystals in 10 mM DFJ solution for 10 min at 293 K. Apo form E566A crystals were obtained using a reservoir solution containing 0.1 M Tris-HCl (pH 7.5), 10% (v/v) isopropyl alcohol, and 15% (w/v) PEG 4000. The E566A-2’FL complex crystals were prepared by soaking the crystals of apo E566A with 10 mM 2’FL for 10 min at 293 K. The crystals of D766A in complexes with l-fucose and lactose were obtained by co-crystallization using a reservoir solution containing 0.1 M MES-Na (pH 6.0), 15% (w/v) PEG monomethyl ether 2000, and 10 mM 2’FL.

Data Collection, Structure Determination, and Refinement—Synchrotron data were collected at beamlines BL-6A and AR-NW12A at Photon Factory (Tsukuba, Japan), BL41XU at SPring-8 (Harima, Japan) and BL9-2 at the Stanford Synchrotron Radiation Laboratory. All data sets were processed and scaled using the HKL2000 program package (29). Phase determination of the selenium-substituted apo form was done by the multiple wavelength anomalous dispersion method using the SOLVE program (30) at 3.0 Å resolution. The initial automatic model was constructed by the RESOLVE program. This model, which contained 1,411 amino acid residues in the asymmetric unit, was used for molecular replacement of the native data set using the Molrep program from the CCP4 program suite (31), and the phases were extended up to 1.12 Å resolution. The structure of the wild type-DFJ complex was determined by the molecular replacement method using Molrep using a refined model of the wild-type apo form as a search model. The structures of the E566A-2’FL and D766A-l-fucose-lactose complexes were determined by the molecular replacement method using the DFJ complex structure as a search model. The automatic model construction for all models was performed using the ARP/wARP program (32). Further model reconstructions were performed manually using Xfit from XtalView (33). Crystallographic refinement was carried out using REFMAC from the CCP4 suite. Coordinates for the DFJ molecule was obtained by modifying the structure of α-L-fucose (Protein Data Bank code 7ABP), and the geometry was optimized using Monomer Library Sketcher from CCP4 prior to incorporation into models. Coordinates for 2’FL, α- and β-l-fucose, and lactose were obtained from Protein Data Bank codes 1GZ9, 1OFZ, and 1IS3, respectively. The qualities of the protein models were assessed with the structure validation program PROCHECK (34). All models have good stereochemistry, and no residues were in the disallowed regions of the Ramachandran plot. Data collection and refinement statistics are summarized in Table 1. Figures were drawn with the programs Molscript (35), Raster3D (36), PyMOL (by W. L. DeLano; available on the World Wide Web at www.pymol.org), and GRASP (38).

Structural Comparison among Related Proteins—The LSQKAB in CCP4 was used to analyze conformational differences among Fuc domain structures. Structural homologs of the Fuc domain were retrieved from the DALI server (39), and the SSM (Secondary Structure Matching) program was used for the structural alignment among these proteins (40).

RESULTS

Overall Structure—We solved the crystal structures of the AfcA Fuc domain in the absence of ligand and in the presence of an inhibitor, substrate, and reaction products. The structure of the unliganded (apo) form was determined at 1.12 Å resolution. The Fuc domain consists of four regions: an N-terminal β region (residues 9–79 and 134–293), a helical linker region (residues 294–386), a helical barrel domain (residues 80–133 and 387–778), and a C-terminal β region (residues 779–896) (Fig. 1a). We then used the DALI server to identify proteins structurally similar to the AfcA Fuc domain, and maltose phos-
Crystal Structure of AfcA Catalytic Domain

phorylase from Lactobacillus brevis (GH family 65) (41) and chitobiase phosphorylase from Vibrio proteolyticus (GH family 94) (42) were highly structurally related, with root mean square (r.m.s.) deviation values of 3.1 and 3.6 Å for the 574 and 504 corresponding Ca atoms, respectively. The high degree of structural similarity was present despite low sequence identities between proteins (maltose phosphorylase, 12%; chitobiase phosphorylase, 11%). The catalytic active sites of both of these phosphorylases reside within the helical barrel domains.

The N-terminal β region of the Fuc domain is composed of 16 antiparallel strands arranged in a super β-sandwich, and it is connected to the four α helices of the linker region. This N-terminal β region has a low degree of structural similarity to Thermaactinomyces vulgaris R-47 α-amylase II (Protein Data Bank code 1BVZ) (43), T. maritima maltosyltransferase (Protein Data Bank code 1GJU) (44), and bovine lysosomal α-mannosidase (Protein Data Bank code 107D) (45), but the function of the N-terminal region remains unclear.

The central helical barrel domain is composed of an (α/α)6 barrel fold domain, and a DALI search identified structural similarities with the catalytic domains of inverting glucoamylases of GH family 15, such as glucoamylase from Thermoanaerobacterium thermosaccharolyticum (Protein Data Bank code 1LF6) (46), glucodextranase from Arthrobacter globiformis (Protein Data Bank code 1UG9) (47), and glucoamylase from Saccharomycopsis fibuligera (Protein Data Bank code 1AUX) (47). The r.m.s. deviation values for the 120-, 118-, and 233-equivalent Cα Data Bank code 1AYX) (47). The r.m.s. deviation values for the 120-, 118-, and 233-equivalent Cα positions of these proteins are 2.8, 2.9, and 3.2 Å, respectively, but they only share 13–20% amino acid sequence identity with the helical barrel domain of the AfcA Fuc domain. The (α/α)6 helical fold is also present in GH65 and GH94 family enzymes. The molecular surface of the central helical barrel domain is characterized by a deep, negatively charged pocket (Fig. 1b), and this region probably represents the substrate-binding pocket. This conclusion is further supported by the solved structures of the AfcA Fuc domain with an inhibitor, substrate, and products (see below). The C-terminal β region forms a two-layered jelly roll fold with a high degree of similarity to the C-terminal domains of maltose phosphorylase and chitobiase phosphorylase, but the function of this region remains unclear.

The dispersive difference Fourier map calculated from the data set at a wavelength of 1.7000 Å (native 2 in Table 1) revealed one strong peak density (above 10 σ) buried within the protein, representing a coordinated metal ion. This metal ion bridges the N-terminal β region, helical linker, and helical barrel domains (supplemental Fig. 1), and three main chain carboxyl groups (from Gly304, Ser385, and Leu392), the oxygen atom of one carbohydrate from Glu376, and one water molecule are the coordinating ligands for this metal ion. We performed inductively coupled plasma emission spectroscopy on the AfcA Fuc domain in solution, and calcium was identified. The ratio of calcium ion to protein molecule was calculated as 0.93 ± 0.1 (data not shown), strongly suggesting that all AfcA Fuc domains bind a single calcium ion.

Inhibitor Complex—DFJ, a nonhydrolyzed L-fucose analogue, is a potent competitive inhibitor of mammalian GH29 α-L-fucosidase (49), and DFJ is also an effective competitive inhibitor for the AfcA Fuc domain, with a Ki value comparable to the Km value for 2-FL. To determine the DFJ binding site, we solved the crystal structure of the AfcA Fuc domain in complex with DFJ at 2.1 Å resolution. The r.m.s. deviation value of the Ca atoms between the apo form and the DFJ-bound complex structure was 0.61 Å, but there was a notable difference between these structures in the central helical barrel domain (Fig. 2a). Three loop regions in the helical barrel domain (Met345–Asn375, Leu638–Tyr698, and Cys711–Thr729) are shifted by 2.1 and 1.7 Å, respectively, and the side chain of Arg677 is rotated by 50° compared with the apo form. The NH2 atom of Arg677 and NE2 of His678 in the long loop (Leu638–Tyr698) are shifted by 2.1 and 1.7 Å, respectively, and the side chain of Arg677 is rotated by 50° compared with the apo form.

The structure of the AfcA Fuc domain in complex with DFJ provided valuable insight into the amino acid side chains relevant for its enzymatic activity. Based on this structure, we mutated several amino acid residues surrounding the bound DFJ molecule, and we determined the enzymatic parameters of the wild-type and mutant proteins (Table 2). Mutations in the conserved acidic residues in GH95 (i.e. E566A and D766A) led to 36,000- and 2,900-fold decreases in the kcat values relative to wild-type AfcA Fuc domain, but the Km values of these mutants were

3 A. Tsuchiya, M. Nagae, S. Wakatsuki, R. Kato, K. Takane, and K. Yamamoto, manuscript in preparation.
Circular dichroism spectra of all mutants were identical with these residues are critically important for fucosidase activity.

Mutation of both Asn421 and Asn423 (N421G/N423G) compared with wild type. Not surprisingly, these data suggest that the conserved acidic amino acids that interact with DFJ are more functionally relevant than a nonconserved acidic residue in the same area. Interestingly, when conserved asparagine residues in the DFJ-interacting region were singly mutated (N421A, N421G, N423G, and N423D), the kcat values of the mutant enzymes decreased drastically compared with wild-type enzyme, but the Km values for these mutants were comparable with wild type.

We next performed chemical rescue experiments to further examine the general acid/base catalysts involved in the reaction. When the residue that acts a general base in a glycosidase is mutated, the addition of an exogenous nucleophilic anionic compound can bind the active site and promote enzymatic activity (51). We used sodium azide as the rescuing nucleophilic anionic compound can bind the active site and promote enzymatic activity (51). We used sodium azide as the rescuing nucleophilic anionic compound can bind the active site and promote enzymatic activity (51). We used sodium azide as the rescuing nucleophilic anion and examined the ability of the mutant enzymes decreased drastically with wild-type enzyme, but the Km values for these mutants were comparable with wild type.

Mutation of both Asn421 and Asn423 (N421G/N423G) completely abolished the enzymatic activity, indicating that these residues are critically important for fucosidase activity. Circular dichroism spectra of all mutants were identical with that of wild-type AfC, indicating that the global secondary structure of the mutant proteins remained intact (data not shown).

Crystal Structure of AfC Catalytic Domain

Table 1

| Data set | Native | Native 2 | Peak | Edge | MAD |
|----------|--------|----------|------|------|-----|
| Data collection statistics of native, native 2, and multiple wavelength anomalous dispersion data sets |
| Space group | P21 | P21 | P21 | P21 |
| Unit cell (Å) | a = 90.2 | b = 112.0 | c = 98.3 | β = 94.7 |
| | b = 90.3 | b = 111.5 | c = 98.5 | β = 95.9 |
| Beam line | AR-NW12A | AR-NW12A | 9.2-2 | 9.2-2 |
| Wavelength (Å) | 1.0000 | 1.0000 | 0.9792 | 0.97936 |
| Resolution (Å) | 98.1-1.12 | 50-2.10 | 128.0-1.90 | 171.7-1.80 |
| Total reflections | 1,062,176 | 3,065,825 | 158,931 | 159,104 |
| Unique reflections | 677,062 | 200,760 | 41,046 | 41,058 |
| Rmerge (%) | 11.2 (33.7) | 9.5 (49.3) | 15.9 (38.3) | 14.2 (7.7) |
| Bond angle (degrees) | 112.1 | 111.5 | 98.5 | 98.3 |
| Bond angle (degrees) | 98.4 | 97.9 | 97.9 | 97.9 |
| Bond length (Å) | 0.016 | 0.015 | 0.014 | 0.014 |
| Bond length (Å) | 0.011 | 0.010 | 0.010 | 0.010 |
| r.m.s. deviations | 1.522 | 1.478 | 1.461 | 1.461 |
| r.m.s. deviations | 1.497 | 1.488 | 1.484 | 1.484 |
| Refinement statistics of native (apo form) |
| Resolution range (Å) | 98.1–1.12 | 50-2.10 | 128.0–1.90 | 171.7–1.80 |
| Total reflections | 1,062,176 | 3,065,825 | 158,931 | 159,104 |
| Unique reflections | 677,062 | 200,760 | 41,046 | 41,058 |
| Rmerge (%) | 11.2 (33.7) | 9.5 (49.3) | 15.9 (38.3) | 14.2 (7.7) |
| Bond angle (degrees) | 112.1 | 111.5 | 98.5 | 98.3 |
| Bond angle (degrees) | 98.4 | 97.9 | 97.9 | 97.9 |
| Bond length (Å) | 0.016 | 0.015 | 0.014 | 0.014 |
| Bond length (Å) | 0.011 | 0.010 | 0.010 | 0.010 |
| r.m.s. deviations | 1.522 | 1.478 | 1.461 | 1.461 |
| r.m.s. deviations | 1.497 | 1.488 | 1.484 | 1.484 |
| Crystal Structure of AfC Catalytic Domain

Data collection and refinement statistics

Table 1

| Data set | Native | Native 2 | Peak | Edge | MAD |
|----------|--------|----------|------|------|-----|
| Data collection statistics of DFJ complex, E566A-2’FL complex, and D766A-fucose-lactose complex |
| Space group | P21 | P21 | P21 |
| Unit cell (Å) | a = 90.7 | b = 111.7 | c = 98.3 | β = 95.3 |
| | a = 88.4 | b = 72.7 | c = 129.2 | β = 96.7 |
| Beam line | BL41XU | BL-6A | AR-NW12A |
| Wavelength (Å) | 1.0000 | 1.0000 | 1.0000 |
| Resolution (Å) | 98.1-2.10 (2.18-2.10) | 128.0-1.90 (1.97-1.90) | 171.7-1.80 (1.86-1.80) |
| Total reflections | 814,509 | 950,638 | 864,168 |
| Unique reflections | 113,829 | 126,307 | 71,461 |
| Completeness (%) | 100 (99.8) | 97.5 (95.3) | 96.8 (96.7) |
| Rmerge (%) | 11.2 (33.7) | 9.5 (49.3) | 15.9 (38.3) |
| Bond angle (degrees) | 19.7 (5.3) | 25.3 (3.8) | 14.2 (7.0) |
| Refinement statistics of DFJ complex, E566A-2’FL complex, and D766A-fucose-lactose complex |
| Resolution range (Å) | 98.1–2.10 | 128.0–1.89 | 171.5–1.80 |
| No. of reflections | 108,104 | 119,917 | 67,841 |
| Rwork (%) | 15.1 | 16.1 | 18.0 |
| Rfree (%) | 19.5 | 20.9 | 21.8 |
| Bond angle (degrees) | 0.016 | 0.015 | 0.014 |
| Bond angle (degrees) | 1.522 | 1.478 | 1.461 |
| r.m.s. deviations | 1.522 | 1.478 | 1.461 |
| r.m.s. deviations | 1.522 | 1.478 | 1.461 |

* Values in parentheses are for the highest resolution shell.

Comparable with the wild-type protein. In contrast, mutation of a nonconserved acidic residue (E485A) caused a 190-fold increase in Km, but only an 11-fold decrease in kcat compared with the wild-type protein. Not surprisingly, these data suggest that the conserved acidic amino acids that interact with DFJ are more functionally relevant than a nonconserved acidic amino acid in the same area. Interestingly, when conserved asparagine residues in the DFJ-interacting region were singly mutated (N421A, N421G, N423G, and N423D), the kcat values of the mutant enzymes decreased drastically compared with wild-type enzyme, but the Km values for these mutants were comparable with wild type.
FIGURE 2. Crystal structure of the Fuc domain in complex with DFJ. a, superposition of the apo form and DFJ-bound complex structures. The main chains of the apo form (pink) and DFJ-bound complex (cyan) are depicted by a wire model. Three acidic residues within the central cavity (Glu485, Glu566, and Asp766) are shown in ball-and-stick models. Three loop regions (Met545–Asn575, Leu638–Tyr698, and Cys711–Thr720) of two structures are highlighted by red (apo form) and blue (DFJ complex), respectively. b, close up view of the inhibitor binding site. The DFJ molecule and the amino acid residues found in the vicinity of DFJ are shown in rod models. The carbon, oxygen, and nitrogen atoms are shown in white, red, and blue, respectively. The water molecule is shown as a red sphere. Hydrogen bonds are depicted by red dotted lines. 2Fo – Fc electron density map of DFJ molecule contoured at the 1.0 σ level is shown in blue mesh. c, structural comparison between the apo form and DFJ-bound complex of the Fuc domain. The main chains of the Fuc domains are shown in wire models colored green (apo form) and cyan (DFJ-bound complex). In the DFJ-bound complex, the DFJ molecule and the amino acid residues discussed throughout are shown in cyan rod models. The corresponding amino acid residues in the apo form are colored green. The distances (in Å) between the C1 atom of the DFJ molecule, Asn421, Glu566, and Asp766 are indicated.
Substrate Complex—Mutational analyses suggested that the conserved Glu\textsuperscript{566} and Asp\textsuperscript{766} residues play critical roles in the enzymatic mechanism of AfcA Fuc domain, and we next solved the crystal structure of the E566A mutant in complex with the substrate 2'FL at 1.9 Å resolution to better determine the molecular role of Glu\textsuperscript{566} in the hydrolysis. The obtained crystal contains two protein molecules in the asymmetric unit, and one contained a clear continuous electron density for the trisaccharide within the active site (molecule B) (Fig. 4a). The other molecule (molecule A) contained only an ambiguous density map at the active site. Additionally, there were considerable conformational differences between these two molecules, with an r.m.s. deviation value of 0.63 Å. The r.m.s. deviation value of all of the Cα atoms between molecule A and the apo form was 0.35 Å, whereas the corresponding value for molecule B was 0.62 Å. In contrast, molecule B was highly structurally similar to the DFJ-bound structure with an r.m.s. deviation value of 0.29 Å. In the crystal packing, the loop region (Leu\textsubscript{638}–Tyr\textsubscript{698}) of molecule A interacts tightly with the symmetry-related molecule through hydrogen bonds. This packing artifact prevents the proper conformational change required for substrate binding in molecule A. From these observations, we conclude that molecule B represents the proper substrate-bound conformation, and we describe the structural characteristics of molecule B below.

The l-fucose moiety of 2'FL adopts a chair \textsuperscript{1}C\textsubscript{4} conforma (Fig. 4b). Molecular recognition of 2'FL by the AfcA Fuc domain is primarily mediated by interactions with the l-fucose and galactose moieties. The mechanism of AfcA Fuc domain l-fucose recognition is essentially identical to that of DFJ (Figs. 2c and 4b). The O4 atom of galactose forms hydrogen bonds with the OE2 atom of Glu\textsuperscript{485} and the NE2 of His\textsuperscript{419}. Significant increase of the $K_m$ value observed in the E485A mutant (Table 2) reflects the loss of the former hydrogen bond. Moreover, slight decrease of the $k_{cat}$ value might be caused by destabilization at the galactose recognition site. For the glucose moiety, only the O6 hydroxyl group forms a weak hydrogen bond with the NH\textsubscript{2} of Arg\textsuperscript{677} at a distance of 3.2 Å. The $k_{cat}/K_m$ value for 2'FL is 7.7-fold higher than that for 2-fucosylgalactose (2FG), which lacks the glucose moiety at its reducing end (data not shown). The change of the binding energy in the transition state ($\Delta AG = -RT \ln((k_{cat}/K_m)_2FG/(k_{cat}/K_m)_2FL)$) (52) for the two substrates was calculated to be 5 kJ/mol, which agrees with a weak hydrogen bond between the NH\textsubscript{2} atom of Arg\textsuperscript{677} and the O6 atom of the glucose moiety. One water molecule (water\textsubscript{2} in Fig. 4b) forms a hydrogen bond with the OD1 atom of Asn\textsuperscript{423}, and it is located near the anomeric carbon atom of the α1,2 glycosidic linkage.

When the structures of the DFJ-bound Fuc domain and the E566A-2'FL complex were superimposed, the position of the l-fucose moiety at the nonreducing end of 2'FL is almost identical to that of the DFJ molecule (Fig. 4c). With the exception of Asn\textsuperscript{421}, the positions of all of the amino acid side chains located in the central pocket coincide. In the DFJ-bound complex, the ND1 atom of Asn\textsuperscript{421} forms a hydrogen bond with the OE1 atom of Glu\textsuperscript{566}. However, in the substrate-bound complex, Asn\textsuperscript{421} cannot form this hydrogen bond due to the point mutation, and the side chain of Asn\textsuperscript{421} is rotated 160° around χ\textsubscript{1} away from the anomeric C1 atom.

Product Complex—Finally, we solved the crystal structure of the D766A mutant in complex with reaction products at 1.8 Å resolution. This structure was obtained by co-crystallizing the D766A mutant in the presence of 2'FL substrate,
and we hypothesized that the substrate would be cleaved during the period of crystal growth. In the obtained crystal, one protein molecule is in the asymmetric unit, and there are distinct electron densities corresponding to L-fucose and lactose in the substrate binding cavity (Fig. 5a). The electron density map surrounding the L-fucose moiety reflects a mixture of \( \alpha \) - and \( \beta \)-L-fucose molecules (Fig. 5a), and the frequencies of the two anomers were calculated as 30 and 70\% by occupancy refinement, respectively. This result is consistent with the equilibration of the mutarotation of liberated L-fucose in the native environment (12) and suggests that the L-fucose molecule reached equilibration with respect to mutarotation during crystal formation. Both L-fucose anomers are also present in the standard \( \text{C}_4 \) chair conformation. The positions of the \( \alpha \) - and \( \beta \)-L-fucose molecules are identical except for the positions of the hydroxyl groups of the anomeric carbon atoms. The equatorial O1 atom of the \( \beta \)-L-fucose is directly hydrogen-bonded with the OD1 atom of Asn\(^{421} \) at a distance of 2.7 Å (Fig. 5b). In contrast, the distance between the axial O1 atom of \( \alpha \)-L-fucose and the
OD1 atom of Asn421 is 4.3 Å. Since the AfC A Fuc domain is an inverting glycosidase, the initial product of its enzymatic activity should be β-1-L-fucose. Therefore, we will focus our discussion on the behavior of β-1-L-fucose.

The mechanism of AfC A Fuc domain reaction product recognition is virtually identical to that of substrate (Fig. 5b). However, two apparent differences that might be caused by the mutational effects are observed in these two structures. The O4 atom of the β-1-fucose moiety forms water-mediated hydrogen bonds with the NE2 atom of Gln764 and the ND2 atom of Asn423, not Asp766. And Glu566 tightly interacts with the hydroxyl O2 and O3 groups in the galactose moiety through the OE2 and OE1 atoms, respectively. In the substrate-bound complex, the glucose moiety is only recognized through a weak hydrogen bond between the O6 hydroxyl group and the NH2 atom of Arg677. Conversely, the distance between these two atoms in reaction product-bound complex is too great to allow for direct interactions.

The overall structure of the D766A products-bound complex is nearly identical to those of the wild type-DFJ- and E566A-2'FL-bound complexes. The r.m.s. deviation values of all Ca atoms for these structures are 0.34 and 0.39 Å,
Crystal Structure of Afca Catalytic Domain

![Image](49x579 to 408x734)

FIGURE 6. Proposed catalytic reaction mechanism of the Afca fucosidase. Hydrogen bonds are depicted by dotted lines. The directions of nucleophilic attack and proton donation are indicated by blank arrows.

DISCUSSION

We report the first crystal structure of a glycosidase from GH family 95. The overall domain arrangement of the Afca fucosidase catalytic domain (Fuc domain) differs substantially from that of T. maritima ι-1-fucosidase (EC 3.2.1.51) of GH29. This protein adopts a two-domain fold, composed of a catalytic (β/α)6-like domain and a C-terminal β-sandwich domain (25). In contrast, the Afca Fuc domain is extremely structurally similar to bacterial phosphorylases from GH families 65 and 94, which are composed of catalytic domains forming an (α/α)6 barrel consisting of six helical hairpin toroids. The (α/α)6 barrel folding topology of the Fuc domain is also common among invertin glycoamylases of GH family 15 in clan GH-L. Comparisons among the structures of Afca Fuc domains and known structurally related enzymes suggested that the active site of the Afca Fuc domain is located within the pocket in the helical barrel domain, and this was confirmed using structural and functional analyses. Little is known about the identity or nature of the catalytic residues in inverting α-l-fucosidases, but the structural similarities of the central (α/α)6 domain to known enzymes provided valuable insight into the enzymatic mechanism of Afca. In inverting glyclosidases, two carboxyl groups serve as general acid and base catalysts. In known structures, these residues are located 10.5 Å apart on average to allow the substrate and the nucleophilic water molecule to bind between the catalytic residues (9–11). The reaction mechanism of inverting phosphorylases is thought to be similar to that of the inverting glycosidase (42). In these enzymes, however, the nucleophilic phosphate ion is thought to directly attack the substrate without activation by the general base residue. Although the Afca Fuc domain has a low degree of sequence identity with the identified structural homologues, Glu566, which is highly conserved in the GH95 family, is virtually superimposable on the catalytic residues of glucoamylases and inverting phosphorylases belonging to the GH15, GH65, and GH94 families. The corresponding residues in all of these proteins are located in topologically identical loops and are thought to act as general acid catalysts during enzymatic activity (41, 42, 46–48). In contrast, Asp766, which we first assumed to be a base catalyst based on the results of mutational and chemical rescue experiments, is located lateral to the general base catalysts in glucoamylases.

Unlike the structure of the apo form, the three complex structures (DFJ-, substrate-, and products-bound) solved here are highly similar. The r.m.s. deviation values for all Ca atoms among all pairs of these three complexes are within 0.39 Å. Therefore, these structures are essentially identical except for the small structural changes caused by the mutational effects, and they probably represent the “bound form” of the enzyme. Upon ligand binding, the domain orientations of bacterial glucoamylase and phosphorylase change (41, 46), but we did not observe such changes in any of the bound structures compared with the apo form of the Afca Fuc domain. This suggests that the (α/α)6 domain is solely responsible for the catalytic reaction. Although the Afca Fuc domain contains a calcium ion, this metal ion probably stabilizes the protein structure rather than playing an important enzymatic role.

The single displacement mechanism used by inverting glycosidases generally relies on two strategically located carboxyl groups derived from either aspartic acid or glutamic acid (9–11). However, our structural and mutational analyses revealed a remarkable reaction mechanism for Afca (Fig. 6). The position of the substituted alanine in the E566A-substrate complex is adjacent to the ι1,2-glycosidic linkage, making it a suitable proton donor for the ι2 atom to aid the release of products. The enzymatic activity of this mutated protein is greatly impaired, indicating that Glu566 acts as a proton donor to the ether oxygen atom of the scissile glycosidic bond.

A catalytic base is required to activate the water molecule for nucleophilic attack at the anomic center for inverting glycosidases. However, no candidate carboxylic acid capable of activating the catalytic water molecule is found in the vicinity of the ι-face of the l-fucose residue in the Afca Fuc domain. Aspartic acid at position 766 is highly conserved, and it is located at the bottom of the catalytic cavity. Due to its location, it cannot directly access the water molecule by the substrate in the crystal structure, suggesting that it seems
to be an inappropriate candidate for the catalytic base. However, substitution of Asp\textsuperscript{766} with alanine severely impairs AfC\textsubscript{A} enzymatic function. Interestingly, one water molecule forms a hydrogen bond with Asn\textsuperscript{421} and is located at a suitable position for in-line nucleophilic attack of the C1 atom of the \(\text{L-fucose} \) moiety (Figs. 6 and 7). The distance between the anomeric C1 atom and the water molecule is 3.0 Å, and the geometry of this arrangement is exquisitely poised “in line,” with an \( \text{O}_2\text{galactose-C1}_{-\text{L-fucose}}-\text{O}_{\text{water}} \) angle of 166°, as expected for the Michaelis complex formed in inverting \( \alpha \)-glycosidases (Fig. 7). Furthermore, the position of this water molecule is nearly identical to that of the equatorial O1 atom of the \( \beta\text{-l-fucose} \) in the products-bound complex. Although the conformation of Asn\textsuperscript{421} in the E566A-substrate-bound complex differs from its position in the wild type-DFJ-bound complex, it is likely that the introduced mutation disrupts the hydrogen bond linking the ND1 atom of Asn\textsuperscript{421} with the OE1 atom of Glu\textsuperscript{566}. Therefore, it is likely that this water molecule forms hydrogen bonds with both Asn\textsuperscript{421} and Asn\textsuperscript{423} in wild-type enzyme (Fig. 4c), giving putative angles for \( \text{C}_\beta\text{Asn}_{\text{421}}-\text{O}_\delta\text{Asn}_{\text{421}}-\text{O}_{\text{water}} \) and \( \text{C}_\beta\text{Asn}_{\text{421}}-\text{O}_\delta\text{Asn}_{\text{421}}-\text{O}_{\text{water}} \) in the wild-type enzyme of 127 and 134°, respectively. These residues could unambiguously define the position and orientation of the water molecule, allowing the lone pair of electrons of the water molecule to face the C1 atom of \( \beta\text{-l-fucose} \) moiety of the substrate. In the D766A-products-bound complex, the OD2 atom of Asn\textsuperscript{421} forms a hydrogen bond with the O1 atom of the \( \beta\text{-l-fucose} \) moiety at a distance of 2.7 Å. These data suggest that two carbonyl groups of Asn\textsuperscript{421} and Asn\textsuperscript{423} play critical roles in withdraw-

ing a proton from the bound water molecule through hydrogen bonds to produce a hydroxide ion, which in turn, acts as a nucleophile to attack the C1 atom of the \( \text{L-fucose} \) moiety.

Since the two asparagine residues (Asn\textsuperscript{421} and Asn\textsuperscript{423}) are predicted to form hydrogen bonds with the corresponding acidic residues (Glu\textsuperscript{566} and Asp\textsuperscript{766}) at distances of 2.9 and 2.6 Å (Fig. 2c), respectively, the enhancement of water nucleophilicity by a carbonyl group of asparagines might be indirectly controlled by these acidic residues, especially Asp\textsuperscript{766}, as deduced from the results of mutational and chemical rescue experiments (Fig. 6). When considered with the structural analyses, one possible explanation for the results of chemical rescue experiments is that the sodium azide molecule in D766A mutant could take the same position as Asp\textsuperscript{766} in native enzyme and enhance the nucleophilicity of Asn\textsuperscript{423}. Interestingly, the N423D mutant exhibited residual enzymatic activity, but this substitution should render residue 423 more suitable as a catalytic base (Table 2). It is possible that the asparagine to aspartic acid substitution led to electrostatic repulsion between N423D and Asp\textsuperscript{766}. Alternatively, an effectual electrostatic relay might be required between Asn\textsuperscript{423} and Asp\textsuperscript{766}. Since a chemical rescue experiment is not always applicable to base/nucleophile mutants, it is not surprising that no rescue reaction was observed in the N421A/G, N423G, or N421G/N423G mutants, whereas these amino acids are important for the enzymatic reaction.

The model that involves such electrostatic relay-mediated activation of the amido group invokes the substrate-assisted catalysis proposed in \( \text{N-acetyl} \beta\text{-hexosaminidase} \) (53, 54), chitinase B (55–57), and \( \text{O-GlCNAC} \) case (58). In the catalytic step of these enzymes, the nucleophilicity of carbonyl oxygen of \( \text{N-acetyl} \) group of a substrate is enhanced and stabilized by the neighboring acidic residue, which makes a hydrogen bond with the imino group of the substrate \( \text{N-acetyl} \) group, and then the activated carbonyl oxygen directly attacks the anomic carbon to form an oxazolinium intermediate (59, 60). Thus, it is not so surprising that an asparagine residue acts as a base if a neighboring carboxylate can act in concert.

Two studies in inverting glycosidases have reported that, in a similar reaction mechanism, an acidic amino acid residue does not function directly as a catalytic acid (61, 62). The catalytic base of celllobiohydrolase Cel6A of GH family 6 is not localized near the anomeric center. In this case, the reaction mechanism is thought to involve the deprotonation of the attacking water via a solvent chain linked to the catalytic base residue. Additionally, inverting enzymes of GH family 48 are thought to use an unusual base catalyst arrangement (63–65). These studies suggest that nonacidic residues can serve as an intermediate in a proton acceptor chain, leading to the activation of a water molecule.

A previous sequence analysis and secondary structure prediction suggested a common evolutionary origin for GH15, -37, -63, -65, -78, -92, -94, and -95 families with conservation of their putative catalytic amino acid residues (37). Two aspartic acids residing in the middle and C-terminal regions of the \( \alpha/\alpha \) domain were thought to be the catalytic amino acids of GH95, and the C-terminal residue corresponds to AfC\textsubscript{A} residue Asp\textsuperscript{766}. However, no residues consistent with
the other predicted catalytic residue is found in the crystal structure, and our structural and biochemical studies clearly identified Glu566, not an aspartic acid, as the general acid catalyst for Af9A. Comparison of structures among families GH15, -65, -94, and -95 reveals conservation of the general acid residues, but Asp766 of Af9A and the corresponding general base residues are shifted. This suggests that the reaction mechanism(s) of GH95 family members differ from those of the GH15, -65, and -94 families. Further supporting this hypothesis is the strong conservation of critical residues lining the enzymatic cavity among GH95 family members (Fig. 8). It is likely that these enzymes also employ the novel reaction mechanism described here.

Acknowledgments—We thank Dr. Stephen G. Withers for useful discussion of the manuscript. We also thank the staff of the beamlines BL9-2 at Stanford Synchrotron Radiation Laboratory (SSRL), BL-6A, and AR-NW12A at KEK-PF and BL41XU at SPring-8 for providing the data collection facilities and for support.
