

Active site complementation and hexameric arrangement in the GH family 29; a structure-function study of α -L-fucosidase isoenzyme 1 from *Paenibacillus thiaminolyticus*

Keywords: crystal structure/ α -L-fucosidase/ glycosidase/ *Paenibacillus thiaminolyticus*

Terézia Kovaľová, Laboratory of Structure and Function of Biomolecules, Institute of Biotechnology Czech Academy of Science, v.v.i., Biocev, Vestec, 252 50, Czech Republic, Fax: +420 325 873 710 Phone: +420 325 873 759 e-mail: kovalovat@ibt.cz

Department of Biochemistry and Microbiology, University of Chemistry and Technology, Prague, Czech Republic Prague, 166 28, Czech Republic.

- Corresponding author (proofs etc.)

Tomáš Koval', Laboratory of Structure and Function of Biomolecules, Institute of Biotechnology Czech Academy of Science, v.v.i., Biocev, Vestec, 252 50, Czech Republic, e-mail: tomas.koval@ibt.cas.cz

Eva Benešová, Department of Biochemistry and Microbiology, University of Chemistry and Technology, Prague, Czech Republic Prague, 166 28, Czech Republic, e-mail: Eva.Benesova@vscht.cz

Patricie Vodičková, Department of Biochemistry and Microbiology, University of Chemistry and Technology, Prague, Czech Republic Prague, 166 28, Czech Republic, e-mail: patricie.vodickova@vscht.cz

Vojtěch Spiwok, Department of Biochemistry and Microbiology, University of Chemistry and Technology, Prague, 166 28, Czech Republic e-mail: spiwokv@vscht.cz.

Petra Lipovová, Department of Biochemistry and Microbiology, University of Chemistry and Technology, Prague, Czech Republic, ICT Prague, Prague, 166 28, Czech Republic,

e-mail: Petra.Lipovova@vscht.cz

Jan Dohnálek, Laboratory of Structure and Function of Biomolecules, Institute of Biotechnology Czech Academy of Science, v.v.i., Biocev, Vestec, 252 50, Czech Republic,

Fax: +420 325 873 710 Phone: +420 325 873 758 e-mail: dohnalek@ibt.cas.cz

- Corresponding author

Running title:

Active site complementation in family GH29

Supplementary data:

1. Size exclusion chromatography absorbance profile
2. Particle size distribution by volume
3. Melting temperature
4. Dependence of relative activity on pH
5. Secondary carbohydrate binding sites
6. Comparison of bound glucose and fucose in the active sites of α -L-f1wt and TMaFuc
7. Calculated binding modes of 3FL and LeX in the active site
8. Summary of active site complementation representatives in GH families
9. Summary of the substrates used in substrate specificity study
10. Thin-layer chromatography plates

Abstract

α -L-Fucosidase isoenzyme 1 from bacterium *Paenibacillus thiaminolyticus* is a member of the glycoside hydrolase family GH29 capable of cleaving L-fucose from non-reducing termini of oligosaccharides and glycoconjugates. Here we present the first crystal structure of this protein revealing a novel quaternary state within this family. The protein is in a unique hexameric assembly revealing the first observed case of active site complementation by a residue from an adjacent monomer in this family. Mutation of the complementing tryptophan residue caused changes in the catalytic properties including a shift of the pH optimum, a change of affinity to artificial chromogenic substrate and a decreased reaction rate for natural substrate. The wild type enzyme was active on most of the tested naturally-occurring oligosaccharides and capable of transglycosylation on a variety of acceptor molecules, including saccharides, alcohols or chromogenic substrates. Mutation of the complementing residue changed neither substrate specificity nor the preference for the type of transglycosylation acceptor molecule, however the yields of the reactions were lower in both cases. Maltose molecules bound to the enzyme in the crystal structure identified surface carbohydrate-binding sites, possibly participating in binding of larger oligosaccharides.

Introduction

α -L-Fucosidases are exo-acting glycoside hydrolases (GH), that catalyse hydrolysis of L-fucosyl residues linked by different types of glycosidic bonds to the non-reducing terminus of oligosaccharides and glycoconjugates. The most frequently occurring types of the linkage are α -1,2 to galactose and α -1,3, α -1,4 and α -1,6 to *N*-acetylglucosamine. α -L-Fucosidases can be found in every kingdom spread through a variety of organisms and tissue types. Their main role lies in nutrient scavenging by polysaccharide degradation and in the ability to modify α -L-fucosylated glycoconjugates, participating in a variety of biological processes (Ashida 2009; Cobucci-Ponzano 2008).

Based on sequence similarity classification by the Carbohydrate-Active-Enzymes database (CAZy) α -L-fucosidases are divided into four families: GH29, GH95, GH141, and GH151 (Lombard et al. 2014; Henrissat 1991). The reaction mechanisms of α -L-fucosidases from GH141 and GH151 are unknown. Family GH141 has only one determined three-dimensional structure, while GH151 has none. Moreover, GH141 also contains xylanases while the other three families only fucosidases. The main difference between families GH29 and GH95 lies in the reaction mechanism. Enzymes belonging to the family GH95 catalyse reactions using a mechanism inverting the anomeric configuration of the released sugar hemiacetal. This process functions via a single displacement mechanism relying on formation of oxocarbenium ion-like transition states. Two amino acid residues play the main role in this mechanism, one acting as a general acid and the other, as a general base. Members of the family GH29 are known for maintaining the anomeric configuration of the released saccharide residue. This is achieved by the double-displacement mechanism of sequential glycosylation and deglycosylation. Two catalytic residues are involved. In the first stage of the reaction, one residue plays the role of the nucleophile, which attacks the anomeric group of the substrate, while the second residue functions as a general acid/base and by protonating

the glycosidic oxygen helps the aglycone moiety to leave. The combined action of these two amino acids leads to the formation of a glycosyl-enzyme intermediate, which is cleaved in the second step of the reaction by a water molecule. Interaction of the intermediate with different acceptor molecules (e.g. alcohols or sugars) leads to transglycosylation (Koshland 1953; McCarter and Withers 1994).

L-Fucose-containing oligosaccharide structures participate in a variety of biological events, such as fertilization or cell adhesion. L-Fucose is a part of human milk oligosaccharides (HMO), which are considered the most important growth factors for bifidobacteria, an important subgroup of commensal human gut bacteria (Bode 2012). α -L-Fucosylated intestinal epithelium supports symbiosis with gut microbiota (Pickard et al. 2014). L-Fucose is also a part of ABO and Lewis blood group antigens (Staudacher et al. 1999; Becker and Lowe 2003; Ma et al. 2006). Changes in composition of oligosaccharides were observed in association with several pathological processes. Increase in α -L-fucosylation of IgG heavy chains was confirmed in rheumatoid arthritis patients (Gornik et al. 1999). A difference in L-fucose to sialic acid ratios was observed in mucins of cystic fibrosis patients compared to control patients (Dische et al. 1959).

α -L-Fucosidase isoenzyme 1 from bacterium *Paenibacillus thiaminolyticus* (*P.thiaminolyticus*) (α -L-f1wt, EC 3.2.1.51) contains 426 amino acids. Its recombinant form produced with a polyhistidine-tag has a molecular weight of 51 kDa and a pH optimum of 8.2 when measured with *p*-nitrophenyl α -L-fucopyranoside (*p*NP α -L-Fuc) as substrate. The protein is a member of the family GH29 (Benešová et al. 2013). Because of the α -L-f1wt transglycosylation abilities, this enzyme is an interesting research target for enzymatic synthesis of α -L-fucosylated compounds. Also, the primary structure of α -L-f1wt shares 26% sequence identity with human tissue α -L-fucosidase (FUCA1), various mutations of which are responsible for the disease called fucosidosis and a 31% identity with human plasma α -L-

fucosidase (FUCA2), a significant factor in the *Helicobacter pylori* infection (Michalski and Klein 1999; Liu et al. 2009). The activity of human α -L-fucosidase in serum is also one of the hepatocellular carcinoma diagnostic markers (Zhang 2015; Miyoshi et al. 2012). Since there are no structures of human fucosidases available, this structure-function analysis of homologous α -L-f1wt can provide some structural data for research on the human enzymes.

Three-dimensional structures of the GH29 family members available today come from enzymes originating from six different organisms (Table I.). Namely there are three different proteins from bacterium *Bacteroides thetaiotaomicron* (BT 2970 - 33%, BT 2192 - 28%, and BT 3798 - 28% sequence identity with α -L-f1wt; calculated using *BLAST*, Altschul et al. 1990). The other structures are of enzymes originating from bacteria *Thermotoga maritima* (29%), *Bacteroides ovatus* (30%), *Bifidobacterium longum* (26%), and one from an unknown bacterial source (32%). The only structure reported for a eukaryotic source is from fungus *Fusarium graminearum* (23%).

In this study, we focus on the structure-function analysis of α -L-fucosidase isoenzyme 1 from bacterium *P. thiaminolyticus*. We present the first three-dimensional structure of the wild type of this enzyme, revealing a novel quaternary structure and active site composition within the family GH29. The observed hexamer represents a functional oligomeric state of the protein, in which the active site is complemented by a residue from an adjacent chain within the hexamer.

Results

Crystal packing and protein fold

Recombinant α -L-f1wt was produced, purified, and used for crystallographic experiments. α -L-f1wt crystallized in the orthorhombic space group $P2_12_12_1$. The structure was determined using molecular replacement and the structure of α -L-fucosidase from *B.*

thetaitotaomicron as a search model (PDB ID: 2WVS, Lammerts van Bueren et al. 2010). The crystal structure contains one hexamer of α -L-f1wt in the asymmetric unit. The amino acid sequence of the enzyme and an overview of the secondary structure elements of one monomer are shown in Figure 1.

A monomer of α -L-f1wt (Figure 2 A) consists of two domains. The N-terminal domain (amino acid residues 1-329) has a fold similar to a $(\beta/\alpha)_8$ barrel. It consists of six β strands and ten α helices. Five short 3_{10} helices are also present. The active site is located inside the β -barrel. The C-terminal domain consists of residues 337-426 in six antiparallel β strands, and a small 3_{10} helix. The monomers of α -L-f1wt assemble into three dimers, which form a hexamer with non-crystallographic 3-fold symmetry. Contacts between two monomers forming a dimer are facilitated mostly by the C-terminal domain of one chain interacting with the N-terminal domain of the opposite monomer (Figure 2 B). An average interface area of this type of contact for all three dimers is 1624 \AA^2 (calculated by PDBePISA, Krissinel and Henrick 2007). The ring-like assembly of the three dimers into the hexamer leads to formation of a channel in the middle (Figure 2 C). The formation of the hexamer is facilitated by interactions of the C-terminal domain of a monomer from one dimer with the N-terminal domain of the monomer from an adjacent dimer. The average interface area of this hexamer-forming type of contact is 530 \AA^2 (calculated by PDBePISA, Krissinel and Henrick 2007). The presence of the hexamer in solution was confirmed by dynamic light scattering (DLS) and size exclusion chromatography (for details see the Supplementary data).

Active site

The active site of α -L-f1wt is located in the N-terminal domain, on the outer side of the hexamer, at a contact of the N-terminal domain with the neighbouring C-terminal domain in

the dimer (Figure 2 B). In the hexamer, two active sites are located opposite each other, around 35 Å apart (Figure 2 C). The active site consists of 15 amino acid residues, based on the distance (4.5 Å) from the bound ligand: Phe21, His23, Glu34, Trp35, His82, His83, Tyr125, Trp184, Asp186, Phe187, Arg218, Glu239, Cys260, Trp267, and Trp392. Out of these residues His23, His82, and Trp184 are conserved within all GH29 members with determined structure. Tyr125 is conserved only in prokaryotic GH29 α -L-fucosidases with solved structure. Asp186, and Glu239 are located at the same position in the structure but at a different position in the sequence within the GH29 family based on the sequence comparison of all α -L-fucosidases with determined structure (Figure 3) using Clustal Omega (Sievers et al. 2011). Among the conserved side chains, there is Asp186, which functions as a catalytic nucleophile and Glu239 which serves as a catalytic acid/base (Figure 4). The catalytic residues were identified based on structural similarities with the structure of α -L-fucosidase from *T. maritima*, which had the residue function confirmed by mutagenesis studies (PDB ID: 1ODU, Sulzenbacher et al. 2004). The active site is complemented by Trp392 from the neighbouring chain within a dimer. Trp392 is a part of a small 3_{10} helix within a loop of the C-terminal domain. The whole loop makes contact with the neighbouring chain, near the active site, using three residues (Trp392, Asn393, and Tyr397). Two of these residues (Trp392, and Asn393) interact with the catalytic acid/base residue Glu239. The distance between Trp392 Cⁿ2 and Glu239 O^ε1 is between 3.3 and 3.8 Å (depending on the chain). Asn393 participates in a hydrogen bond with the main chain carbonyl of Glu239. The mutual orientation of the side chains of Trp392 and Trp35 resembles a carbohydrate binding ‘platform’ (Figure 4 B), the shortest distance between their Cⁿ2 atoms being 7.8 Å. Such platforms formed by aromatic residues are a common part of carbohydrate-binding proteins (Boraston et al. 2004). This is the first documented structure of a GH29 family member with the active site complemented by the second protein chain.

Inside the active site, electron density for a ligand was observed, resembling a saccharide molecule (Figure 4 C). As a first choice a maltose molecule was considered, because it was used as an additive in crystallization. However, the $2mF_o - DF_c$ electron density after refinement did not support the idea of a disaccharide and suggested rather a monosaccharide. This led to consideration of glucose as a possible ligand. Glucose was present as a contaminant in maltose (< 0.5% according to the manufacturer; *i.e.* < 0.25 mM concentration in the crystallization condition). Presence of glucose in the maltose sample was confirmed by mass spectrometry. No saccharides were detected in the protein sample prior crystallisation using mass spectrometry. Glucose molecule was modelled and refined in different conformations. The ¹E conformation appeared as the best fit (Figure 4 D). The comparison of the position of modelled D-glucose and of bound L-fucose in the structure of TMaFuc can be seen in Supplementary data.

A non-Pro *cis* peptide bond occurs between residues Asp186 (catalytic nucleophile) and Phe187, fully supported by electron density. Phe187 together with Trp35 and Phe39 form a carbohydrate binding platform. Presence of this *cis* peptide in this structure enables proper positioning of both residues in the active site. It is a rare type of bond conformation as most *cis* peptides occur between any amino acid residue and Pro. A study conducted on a non-redundant set of 571 protein structure from the PDB revealed that 5.2% of peptide bonds are in *cis* conformation formed between other amino acid and Pro, while the occurrence of a non-Pro *cis* peptide bond in the general case is about 0.03%. One third to one half of non-Pro *cis* peptide bonds are either in carbohydrate-binding or carbohydrate-processing proteins (Jabs et al. 1999). Within the group of structurally characterized α -L-fucosidases we found only one other case in which a Phe residue from the platform of aromatic residues is present right next to the catalytic nucleophile and is also the following residue in the sequence – in α -L-

fucosidase from *F. graminearum* (PDB ID: 4PSP; Cao et al. 2014). In this case, however, the peptide bond is in *trans* conformation and the local structure differs from α -L-f1wt.

Secondary carbohydrate binding sites

Maltose molecules were observed bound in three different sites in the hexamer (Figure 2 C). The first carbohydrate binding site (CBS; number I) is at the interface of the N-terminal domain of chain A and the C-terminal domain of chain E. Here, two molecules of maltose are bound stacked on top of each other. The residues involved in binding of the two maltose molecules are Ser53, Asn55, and Val57 of the N-terminal domain and Asp328, Lys349, Asn351, Lys352, and Glu419 of the C-terminal domain. Two maltose molecules stacked on top of each other are also present at a similar site between chains C and A, and then D and F. There are no maltose molecules bound in the other three identical sites of type I, probably because of the different crystal packing environment. The distance between the nearest active site and site I is around 40 Å, which corresponds to the length of an oligosaccharide about twelve monomers long. Another maltose molecule is also bound to the N-terminal domain of chain E (site II), making contacts with eight residues from this chain (Arg37, Val44, Glu45, Trp131, His132, His144, Phe152, and Arg155) and Asn160 from chain A of the neighbouring α -L-f1wt molecule. The distance between this site and the active site is around 37 Å. It cannot be excluded that the observed carbohydrate binding sites I and II belong to binding sites for larger substrates. The third carbohydrate binding site (site III) is located at the crystal contact of two hexamers, between chains E from one hexamer (residues Asn175-Tyr177, Arg109, and Asp113) and D from another (Gln8, Gln11, His15, and Asp16) and probably helps stabilize the crystal. Detailed views of the individual CBSs are available in Supplementary data.

Catalytic study

A mutation of the active site residue Trp392 to alanine was designed in order to understand its role. The oligomeric state of the protein was not changed by the mutation as confirmed using DLS (for details see the Supplementary data). The mutant (further denoted as α -L-f1W392A) was used alongside with α -L-f1wt in catalytic assays to compare their catalytic abilities. The hydrolytic activity of both proteins was tested using *pNP* α -L-Fuc as substrate. The specific activity of the used protein samples was 275 ± 9 μ mol/min/mg for α -L-f1wt and 190 ± 4 μ mol/min/mg for α -L-f1W392A. The mutant was also used in crystallization experiments, but it did not produce any crystals.

The optimal pH for catalysis was determined using a set of the Britton-Robinson buffers (pH range 1.9-10.2) with 10 mM *pNP* α -L-Fuc as substrate. Optimal pH of α -L-f1wt is about 8.5, while for α -L-f1W392A it lies between 7.0 and 7.5 (see Supplementary data for further details).

Dependence of initial rate on the substrate concentration was tested with *pNP* α -L-Fuc and 2'-fucosyllactose (2FL) as substrates (Figure 5). In the case of 2FL the assay was limited to monitoring of released fucose. The kinetic parameters are calculated using the equation for simultaneous hydrolysis and transglycosylation according to Kawai et al. (2004) and they are summarised in Table II. The simultaneous transglycosylation was confirmed by detection of transglycosylation products in the reaction mixture using mass spectrometry. For *pNP* α -L-Fuc, the mutation caused an increase in K_M , while the maximal rate changed only slightly. When 2FL was hydrolysed, K_M for the mutant-catalysed reaction stayed about the same, while the maximal rate decreased three times compared to the wild type.

The ability to cleave α -L-fucose from different substrates was tested for both proteins. The results for all substrates are summarized in Table III. The progress of the reaction was monitored using thin-layer chromatography (TLC; for details see the Supplementary data).

The mutation negatively affected the rate of reaction, which resulted in lower yields. However, it had no effect on the type of substrate the enzyme can deglycosylate.

The transglycosylation capabilities of α -L-f1wt and α -L-f1W392A were tested using a set of selected acceptor molecules, which included several alcohols, saccharides, amino acids and *p*-nitrophenyl glycopyranosides. The summary of the results for the transglycosylation reaction catalysed by α -L-f1W392A can be found in Table IV. The data for α -L-f1wt were already published by Benešová et al. (2013). The reaction was monitored using TLC (for details see the Supplementary data). The designed mutation did not affect the enzyme's ability to transfer L-fucose to a specific type of the acceptor molecule, it only affected the rate of the reaction catalysed by α -L-f1W392A. Attempts were made to obtain structures of complexes of the enzyme with natural products and acceptors but the experiments did not lead to crystals suitable for x-ray analysis.

Docking of 3FL and LeX into the active site

Possible binding modes of two trisaccharides substrates (3FL and LeX) differing only in the presence of *N*-acetylation were examined using the PLANTS program. The binding pose of 3-fucosyllactose with the best docking score (-89.009) was in a good agreement with the experimentally determined binding poses of glucose (this work) or fucose analogue in the structure of *Bacteroides thetaiotaomicron* α -L-fucosidase (PDB ID: 2WVS, Lammerts van Bueren et al. 2010, for graphical comparison see Supplementary data). The binding pose of Lewis X trisaccharide had a slightly better score (-90.5992) but it was not in agreement with the known fucose binding mode. The docking pose closest to the known fucose binding mode was the pose number 10 with a poor score (-75.1283). The docking results are in agreement with the difference in the activity of α -L-f1wt towards 3-fucosyllactose and Lewis X trisaccharide.

Melting temperature and oligomeric state of α -L-f1wt and α -L-f1W392A

In order to analyse the effect of the mutation on structural stability of α -L-f1wt, the wild-type and the mutant were analysed using DLS and differential scanning fluorimetry (DSF). DLS measurements did not show any change of the oligomeric state of the mutated protein compared to the wild-type. The DSF measurement analysing the melting temperature showed a little difference between the wild-type and the mutant (α -L-f1wt: 59 °C; α -L-f1W392A: 60 °C). The data from both experiments can be found in Supplementary data.

Discussion

Comparison of the oligomeric assembly of the known α -L-fucosidases

To date eight different α -L-fucosidases from family GH29 have their three-dimensional structures recorded in the PDB (Table I). Most of them come from bacteria, except for the fungal α -L-fucosidase from *F. graminearum*. Among these eight proteins, only α -L-fucosidase from *T. maritima* (TMaFuc), is known to occur in a hexameric state in solution (PDB ID: 1ODU, Sulzenbacher et al. 2004). The hexameric assembly of TMaFuc significantly differs from the α -L-f1wt hexamer (Figure 6). For both, TMaFuc and currently reported α -L-f1wt, the presence of the hexameric state in solution was confirmed by size-exclusion chromatography and it probably represents the active form of these enzymes. The way in which these two proteins assemble their monomers into a hexamer differs in several key aspects. Both hexamers assume a ring-like arrangement containing two trimers, stacked on top of each other. However, α -L-f1wt adopts a ring-like shape with two inner channels, one visible in view along the non-crystallographic 3-fold axis and one visible in view along the non-crystallographic 2-fold axis. TMaFuc forms a hexamer, with a depression in the middle of the hexamer in view along the non-crystallographic 3-fold axis. The only channel present in TMaFuc is located between the two stacked trimers, along the non-crystallographic 2-fold

axis. In TMaFuc the main contacts involved in dimer formation are made only by the N-terminal domains, while the C-terminal domains point outwards from the ring, into the surrounding solvent. In contrast, the α -L-f1wt dimer formation is facilitated by contacts of the N-terminal domain of one monomer and the C-terminal domain of the second monomer. Three dimers of α -L-f1wt assemble in a hexamer with a ring-like shape by forming contacts between the C-terminal domain of a monomer in the first dimer and the N-terminal domain of a monomer in the second dimer. The dimers of TMaFuc, however, assemble in a hexamer utilizing only contacts involving their N-terminal domains. The calculated interface area of the TMaFuc dimer is 895 Å², while for the α -L-f1wt dimer it is almost doubled (1624 Å²). The active sites are located in both hexamers on the outer side of the ring. In both cases two active sites from different monomers within a hexamer are grouped together, facing each other. The shortest distance between such two active sites in the TMaFuc hexamer is about 40 Å and in α -L-f1wt around 35 Å.

The α -L-f1wt hexamer has a 32 point symmetry, which was observed also in several other glycosidases, such as in β -galactosidase from *Arthrobacter* sp. C2-2 (Skálová et al. 2005).

Comparison of tertiary structure

While the hexameric assembly of α -L-f1wt and TMaFuc differs, their fold is similar (C^α r.m.s.d. 1.6 Å, calculated by Coot, Emsley et al. 2010). They both contain two domains, the catalytic N-terminal domain with a $(\beta/\alpha)_8$ barrel-like fold and the C-terminal β -sandwich domain (Figure 7). The described structure is quite common in the GH29 family, although there are cases of α -L-fucosidases containing three domains, one N-terminal catalytic domain with a $(\beta/\alpha)_8$ barrel-like fold and two β -sandwich domains (PDB ID: 4ZRX, Unpublished; 4PSP, Cao et al. 2014; 5K9H, Summers et al. 2016).

Active site complementation

The most significant difference in the tertiary structure of the C-terminal domains of α -L-f1wt and TMaFuc is the presence of a 10-residue loop (residues 390-399) with a small 3_{10} helix in α -L-f1wt (Figure 7). This loop contains Trp392, which is a part of the active site of the adjacent monomer, and together with Trp35 is creating a carbohydrate binding motif called Tryptophan box (Yang et al. 2015). No active site complementation is present in any other α -L-fucosidase with known structure. Active site complementation was observed in other GH families (GH2, GH13, GH20, GH31, GH35, GH38, GH42, GH94, GH97, and GH117; summary of structures and references available in Supplementary data). The representatives of each GH family with active site complementation were identified by manual checking of glycosidase hydrolase structures in PDB reported in multimeric assembly. The individual assemblies of the active sites differ in the number of monomers participating in an active site construction and in the types and numbers of residues involved. The most similar case to that of α -L-f1wt can be found in hexameric α -xylosidase from *E. coli* (GH31, PDB ID: 1XSI, Lovering et al. 2005). This enzyme has its active site composed of residues from three different protein chains. The monomer of α -xylosidase has two distinct loops on its N-terminal β -sandwich domain. One of them is a tryptophan-containing loop similar to that of α -L-f1wt (Figure 8). This loop is adding a tryptophan residue (Trp8) to the active site in a similar position with respect to the catalytic residues but in a different orientation than in α -L-f1wt. The second loop (from a chain different from that bearing Trp8) introduces Leu48 and Asp49 into the active site.

Influence of Trp392 on catalytic activity

The effect of Trp392 on catalytic activity was examined using the mutant α -L-f1W392A, natural substrate (2FL) and artificial substrate (*p*NP α -L-Fuc). The results showed differences in the catalytic activity caused by the mutation. In the case of natural 2FL, the maximal rate of the reaction catalysed by the mutant dropped three times, while K_M stayed about the same.

When *pNP* α -L-Fuc was used as substrate, the maximal rate of the catalysed reaction of the mutant changed slightly, while K_M increased ten times. α -L-f1wt catalyses reaction using the double-displacement reaction mechanism with two steps, glycosylation and deglycosylation of the protein. Chromogenic substrates with a good leaving group as an aglycon, such as *pNP* α -L-Fuc undergo rapid glycosylation followed by slow deglycosylation, which is then the rate-limiting step. Mutation W392A affects substrate recognition and the glycosylation step, which influence affinity for *pNP* α -L-Fuc (K_M), but not the rate of the reaction which is determined by the deglycosylation step. The rate-limiting step in the case of a natural substrate such as 2FL is usually the slow glycosylation, which is also affected by the mutation in our case and results in a decrease of the reaction rate. The affinity for the trisaccharide substrate was affected less by the mutation as the K_M stayed about the same. The observed changes in maximal rate and substrate affinity for the natural and artificial substrate demonstrate that Trp392 has an influence on the catalytic activity of α -L-f1wt, which makes this the first observed member of family GH29 with the active site complemented by a residue from the adjacent monomer and accompanied catalysis effects.

Substrate specificity of α -L-f1wt and α -L-f1W392A was tested on a selected group of naturally occurring substrates (Table III). The substrate molecules differed in the number of monosaccharides and the presence of branching in the oligosaccharide. The wild-type protein demonstrated the ability to cleave all common types of the glycosidic bond and almost all tested substrates. The only exception was substrate Lewis X (α -1,3 bond to *N*-acetylglucosamine). The rest of the tested substrates were all hydrolysed with higher or lower yields.

To examine the differences in substrate specificity docking computations with ligands Lewis X and 3FL (differing in presence of one *N*-acetyl group) were performed. The energetically most favourable binding pose of Lewis X trisaccharide was influenced by its *N*-acetyl group

and placed the fucose moiety away from the catalytic residues, which would disable catalysis. Docking of 3-fucosyllactose led to a binding pose consistent with catalysis and with an orientation to the catalytic residues similar to the experimentally observed binding pose of fucose derivative in the structure 2WVS.

The mutation W392A had no effect on the type of the cleaved bond, or on the type of the oligosaccharide, from which L-fucose was cleaved. It only affected the rate of the reaction. The transglycosylation abilities of the mutant were also tested on the selected group of acceptor molecules and compared with the wild type (Table IV). The mutation caused no change in the preference of the type of the acceptor molecule. The only effect of the mutation was a lower yield of the transglycosylation reactions. It is possible that the changes in activity caused by the mutation could be a consequence of the close proximity of the mutated tryptophan residue to the catalytic acid/base. It is likely that Trp392 is involved in the active site pocket formation. The removal of its side chain could result in an active site remodelling which would complicate binding of a substrate in the right orientation for catalysis to occur.

The mutation affected the pH optimum of the protein. It was about 1 point lower (7.0-7.5) than that of the wild type (8.5). This change was probably caused by the close proximity of the mutated tryptophan to the catalytic acid/base residue (Glu239). Its mutation to alanine made the environment less hydrophobic and thus probably lowered the pK_a of the catalytic acid/base residue. It cannot be excluded that the mutation causes some active site remodelling, which would provide another explanation of the observed change. Such remodelling would also affect the pK_a value of the catalytic residue. Efforts to see the changes in a structure of the mutant were unsuccessful.

Effect of mutation on structure stability

The mutation of Trp392 to alanine did not cause any changes to the oligomeric state of the protein as confirmed by DLS. The mutant remained in the hexameric state. The thermal stability of the mutant was almost unaffected, which together with the maintained oligomeric state suggests that this residue has no significant structural role in the stabilization of the protein structure.

Carbohydrate binding module-like domain

The C-terminal domain of α -L-f1wt containing the Tryptophan box-forming residue Trp392 has a fold that resembles carbohydrate-binding modules (CBM), common parts of many different glycoside hydrolases. Our sequence search using BLAST (Altschul et al. 1990) did not find any positive match to any known member of the eighty-three so far described CBM families. A structural search using the C-terminal domain and PDBeFold (Krissinel and Henrick 2004) found the closest CBM by fold (C^α r.m.s.d. 2.25 Å) of endoglucanase D from *Clostridium cellulovorans* (PDB ID: 3NDY, to be published). This domain is classified as CBM2 (CAZy, Lombard et al. 2014). The folds of the C-terminal domain of α -L-f1wt and of CBM2 members differ in the position of loops and the length of the beta sheets. A sequence search with the C-terminal domain of α -L-f1wt using BLAST found over one hundred proteins with sequence identity over 50% and full coverage of the domain; all of them were bacterial α -L-fucosidases, mostly from genera *Paenibacillus* or *Bacillus*. This demonstrates a strong genetic conservation of this domain and together with its participation in binding carbohydrates, it appears that it is a member of a so far undescribed CBM family.

Ligand in the active site

The electron density for a ligand in the active site was interpreted as a glucose molecule in the 1E conformation, which is a higher energy state than the typical 1C_4 conformation. Glucose was reported as an inhibitor of α -L-fucosidase from *Chamelea gallina* (Reglero and Cabezas

1976) and of human α -L-fucosidase (21% inhibition at 6 mM; Avila and Convit 1974) There are known structures of α -L-fucosidase BT 2970 from *B. thetaiotaomicron* with an *fuco*-configured inhibitor molecule ((2*S*,3*S*,4*R*,5*S*)-2-(1*H*-benzol[d]imidazol-2-yl)-5-methylpyrrolidine-3,4-diol) bound in the active site in the ³E conformation (Wright et al. 2013). It is known, that during catalysis the geometry of a saccharide substrate in an active site of a glycoside hydrolase is distorted (Mayes et al. 2014). The unusual conformation adopted by the glucose molecule in our case may be a result of the natural distortion of the pyranose ring that occurs within the active site during the catalysis (Mayes et al. 2014; Davies et al. 2012).

Conclusion

The three-dimensional structure of α -L-fucosidase isoenzyme 1 from *P. thiaminolyticus* solved using x-ray crystallography reveals a new quaternary structure and active site composition within the family GH29. The protein is in a new type of a hexameric assembly, which represents the active form of the enzyme. The structure also shows a unique active site complementation by a tryptophan residue (Trp392) from the adjacent monomer within the hexamer. This residue is a part of the carbohydrate-binding site. Trp392 is important for the binding and cleavage of different substrates and its presence influences the pH optimum, affinity to some substrates and the reaction rate. Several carbohydrate binding sites were identified in the structure. Besides the active site, there are also two different binding sites containing co-crystallized maltose in the structure. The additional sites could form parts of more extensive carbohydrate binding sites for larger substrates.

Materials and methods

Expression and purification of α -L-f1wt

Recombinant protein α -L-fucosidase isoenzyme 1 from *P. thiaminolyticus* in fusion with polyhistidine-tag was expressed using competent cells *E. coli* BL21 (DE3) and expression vector pET16b- α LF1 prepared previously by Benešová et al. (2013). Transformed cells were cultivated in 600 mL of Luria-Bertani (LB) medium with 0.1 mg/mL ampicillin at 120 RPM and 37 °C overnight. Cells were harvested by centrifugation at 3000 RPM at 4 °C for 7 min. Cells were disintegrated by incubation with hen egg white lysozyme (Fluka, USA) at 25 °C for 30 min followed by incubation with sodium deoxycholate (0.1% in the reaction mixture) at 4 °C for 30 min. Afterwards, DNase I (New England Biolabs, Ipswich, MA, USA) was added (4 U), and cells were incubated at room temperature for 15 min. Disintegration was concluded by sonication (20 W, 6 \times 30 s) on ice using an ultrasonic liquid processor Sonicator 3000 (Misonix, Inc., Farmingdale, NY, USA). The resulting lysate was centrifuged at 20 000 g for 20 min and the supernatant was used in the purification step. α -L-f1wt was purified to homogeneity using affinity chromatography with Ni-NTA agarose and transferred into 50 mM phosphate buffer, pH 7.5 using a PD10 desalting column (GE Healthcare, UK) according to Benešová et al. (2013). The desalted sample was further purified using size exclusion chromatography with an ÄKTA Prime (Amersham Bioscience, USA) and a HiLoad TM Superdex 200 prep grade column (GE Healthcare, UK). The elution buffer contained 50 mM phosphate buffer, pH 7.5. The purity of the isolated protein was assessed by SDS PAGE using gel composed of 5% stacking and 10% running gel.

Expression of α -L-f1W392A

Mutant with Trp392 mutated to Ala was prepared using two primers with the desired mutation (forward 5'-AG ACG CCGGCC AAT GCG GTC GAC TAC CC- 3'; reverse: 5'-AC CGC ATTGGC CGG CGT CTG AAT GTT GACC- 3', Generi Biotech, Hradec Králové, Czech Republic). PCR mixture was prepared according to manufacturer's instructions. It contained 120 ng of template plasmid (pET16b- α LF1) DNA and 0.02 U/ μ L of KOD Hot Start DNA

Polymerase (Novagen, Madison, WI, USA). PCR started with initial denaturation at 95 °C for 2 min and continued with the program consisting of 10 cycles of denaturation at 95 °C for 20 s, annealing at 70 °C for 10 s, extension at 70 °C for 4 min, and 20 cycles of denaturation at 95 °C for 20 s, annealing at 65 °C for 10 s, extension at 70 °C for 4 min, and final extension at 70 °C for 10 min. The resulting PCR product was incubated with *Dpn1* (New England Biolabs, Ipswich, MA, USA) at 37 °C overnight to cleave all template DNA and afterward purified using a QIAquick PCR purification kit (QIAGEN, Hilden, Germany). The Gibson Assembly Master Mix (New England Biolabs, Ipswich, MA, USA) was used as described by the manufacturer to reconnect the pET16b- α LF1W392A plasmid which was then used to transform *E. coli* DH5 α cells. 7 μ L of reaction mixture was added to 100 μ L of freshly thawed competent cells, incubated on ice for 20 min, then heated up to 42 °C for 90 s. After heating, 400 μ L of LB medium was added and cells were incubated at 37 °C for 1 h. 200 μ L of incubated cell mixture was spread on LB agar plates containing 0.1 mg/mL of ampicillin and left at 37 °C overnight. 14 grown colonies were separately transferred into 4 mL of LB medium with 0.1 mg/mL ampicillin and incubated at 37 °C and at 120 RPM overnight. Cells were harvested and plasmid DNA was isolated by alkaline lysis (Birnboim and Doly 1979). Isolated plasmids were analysed by agarose electrophoresis (1% gel), and their concentration was measured using a NanoDrop (Thermo Fisher Scientific, USA). The presence of the mutation and accuracy of the vector were confirmed by sequencing performed by GATC Biotech (Germany).

Plasmid pET16b- α LF1W392A was used to transform *E. coli* BL 21 (DE3) cells as described previously. Transformed cells were placed on LB plate supplemented with 0.1 mg/mL ampicillin and incubated at 37 °C overnight. All grown colonies were transferred into 600 mL of LB medium with 0.1 mg/mL ampicillin and incubated at 37 °C and at 130 RPM overnight. Cells were harvested by centrifugation at 3000 g for 15 min, and 4 °C and resuspended in

25 mM EPPS buffer, pH 8. Protein α -L-f1W392A was isolated and purified to homogeneity using the same protocol as the wild type.

Dynamic light scattering

The oligomeric state of the protein α -L-f1wt was examined using DLS with a Zetasizer Nano ZS90 (Malvern Instruments, UK) and a 25 μ L quartz cuvette. The measurement was conducted at 18 °C with 1 mg/mL protein sample.

Crystallization of α -L-f1wt

Recombinant α -L-f1wt was concentrated using 10 kDa Microsep™ Advance Centrifugal Device (PALL Port Washington, NY, USA) and used for crystallization. The final concentration of the protein was determined spectrophotometrically at 280 nm 15 mg/mL. Initial crystallization screening was done using the Index crystallization screen (Hampton research, USA) and the hanging drop vapour diffusion method at 18 °C. The ratio of the protein to reservoir drop volume was 1:1 (0.5 μ L + 0.5 μ L). Several successful hits were further optimized. The best crystals were grown using the sitting drop vapour diffusion method and Crystal Quick 96 Well Plates (Greiner, USA) sealed with Clear Seal Film (Hampton Research, USA) at 18 °C. The reservoir solution contained 25% (w/v) PEG 3350, 0.1 M BIS-TRIS buffer, pH 6.5, and 0.2 M ammonium acetate. The volume of the reservoir was 70 μ L. 0.5 μ L of reservoir solution diluted to 80% of the original concentration and containing 50 mM maltose was added to 0.5 μ L of the protein solution.

Data collection and structure determination

Crystals were cryo-protected in perfluoropolyether oil and vitrified in liquid nitrogen using round LithoLoops (Molecular Dimensions, USA). Diffraction data were collected at the synchrotron radiation source DESY: PETRA III, (Hamburg), at the beamline P13, using a

Maatel MD2 micro-diffractometer with a mini kappa goniometer and a Dectris Pilatus 6M-F detector at 100 K. Data were indexed, integrated and scaled using XDS (Kabsch 2010). Merging was done using AIMLESS (Evans et al. 2013) from the CCP4 program suite (Winn et al. 2011). The phase problem was solved by molecular replacement using MOLREP (Vagin and Teplyakov 2010) and the structure of α -L-fucosidase from *B. thetaiotaomicron* (PDB ID 2WVS, Lammerts van Bueren et al. 2010). The structure was refined using REFMAC5 (Murshudov et al. 2011), and model building was carried out using COOT (Emsley et al. 2010). Data collection and refinement statistics are shown in Table V.

Bioinformatics

Sequence homologues of α -L-f1wt within the databases were searched for using BLAST (Altschul et al. 1990). Sequence alignments were calculated using Clustal Omega (Sievers et al. 2011). $2F_o - F_c$ composite omit map was calculated using PHENIX (Adams et al. 2010). Structural homologues were searched for using PDBeFold (Krissinel and Henrick 2004). The average interface area of the hexamer-forming contacts was calculated by PDBePISA, (Krissinel and Henrick 2007). Secondary structure elements of α -L-f1wt were assigned based on chain A using PDBeFold (Krissinel and Henrick 2004).

Docking experiments

Docking was performed using the PLANTS program (Korb et al. 2009) with the scoring function CHEMPLP and search speed set to 1. The centre of the binding site was set to [7.3, 27.3, 56.1 Å] and its diameter was set to 14 Å. All non-protein atoms were removed prior docking and hydrogens were added using UCSF Chimera (Pettersen et al. 2004). Structures of 3-fucosyllactose and Lewis X trisaccharide were built using the Glycam server (Kirschner et al. 2008).

Activity assay

The glycolytic activity of α -L-f1wt and α -L-f1W392A was tested using *p*NP α -L-Fuc (Carbosynth, UK), where released *p*-nitrophenol was measured spectrophotometrically at 405 nm. The final volume of the reaction mixture was 110 μ L and the mixture contained 10 mM *p*NP α -L-Fuc, in 25 mM EPPS buffer, pH 8, and 0.0174 μ g of α -L-f1wt or 0.0198 μ g of α -L-f1W392A. The reaction was started by addition of enzyme and stopped by addition of 100 μ L of 10% Na₂CO₃. Reaction time was 10 min at 37 °C. 1 U is the amount of the enzyme capable of releasing 1 μ mol of *p*-nitrophenol per 1 minute.

Kinetic parameters of α -L-f1wt and α -L-f1W392A were determined using *p*NP α -L-Fuc and 2'-fucosyllactose (2FL, Carbosynth, Compton, UK) as substrates. Concentration of *p*NP α -L-Fuc was in the range 0.25-25 mM (diluted in 25 mM EPPS buffer, pH 8). The range of 2FL concentration was 5-250 mM (diluted in water). All reactions were carried out at 37 °C for 10 minutes in 25 mM EPPS buffer, pH 8. In the case of *p*NP α -L-Fuc, released *p*-nitrophenol was measured spectrophotometrically at 405 nm. For detection of released L-fucose, an L-fucose assay kit was used (Megazyme, Bray, Ireland). Kinetic parameters were calculated with an account of simultaneously ongoing transglycosylation using the equation according to Kawai et al. 2004. For calculating the K_M and V_{lim} the minimal kinetic scheme for substrate hydrolysis and transglycosylation was used (Figure 9; Nguyen et al. 2014; Kawai et al. 2004).

Released lactose can also function as acceptor of fucose in transglycosylation. As the reaction is monitored only for a short period, in which the concentration of 2FL is much higher than that of released lactose, the contribution of lactose as acceptor is minimal.

The pH optima of α -L-f1wt and α -L-f1W392A for the hydrolysis of *p*NP α -L-Fuc were determined by measuring their activity using the Britton-Robinson set of buffers (pH range 1.9-10.2). The reaction mixture contained 50 μ L of 10 mM *p*NP α -L-Fuc (diluted in water) and 50 μ L of buffer. The reaction was started by addition of 10 μ L of diluted enzyme. The concentration of α -L-f1wt was 1.8 μ g/mL, and of α -L-f1W392A 2.8 μ g/mL. Reactions were

carried out at 37 °C for 10 min and were stopped by the addition of 100 μ L of 10% Na₂CO₃. Released *p*-nitrophenol was measured spectrophotometrically at 405 nm.

Substrate specificity assay

The ability of α -L-f1wt and α -L-f1W392A to cleave L-fucose from various oligosaccharides was tested using nine L-fucose containing substrates, namely 2'-fucosyllactose, 3-fucosyllactose, Lewis A trisaccharide, Lewis X trisaccharide, Globo-H, blood group H disaccharide, 2-acetamido-2-deoxy-3-*O*-(α -L-fucopyranosyl)-D-glucopyranose, 2-acetamido-2-deoxy-6-*O*-(α -L-fucopyranosyl)-D-glucopyranose, 2-acetamido-2-deoxy-4-*O*-(α -L-fucopyranosyl)-D-glucopyranose (Carbosynth, Compton, UK). Both enzymes were diluted to 0.3 mg/mL in 25 mM EPPS buffer, pH 8 and mixed with 20 mM substrate in ratio 1:1. The final volume of the reaction mixture was 6 μ L. Reaction time was 3 h and reaction temperature was 40 °C. The release of L-fucose was detected using TLC.

Transglycosylation assay

Transglycosylation abilities of α -L-f1wt and α -L-f1W392A were compared using twenty different acceptor molecules selected from alcohols, saccharides, amino acids and *p*-nitrophenylated saccharides as described previously (Benešová et al. 2013). Reaction mixtures were analysed using TLC. Saccharide acceptors were tested as follows: 50 mM saccharide (concentration in reaction mixture) dissolved in 25 mM EPPS buffer, pH 8 and 50 mM *p*NP α -L-Fuc (dissolved in DMF) was incubated with 0.06 mg/mL of protein at 45 °C for 3 h. For testing amino acids as acceptors a 10 mM solution of an amino acid (concentration in reaction mixture) dissolved in 25 mM EPPS buffer, pH 8 and 50 mM *p*NP α -L-Fuc (dissolved in DMF, concentration in reaction mixture) were incubated with protein (0.06 mg/mL in reaction mixture) at 45 °C for 3 h. Alcohol acceptors were tested as follows: 30% methanol solution (concentration in reaction mixture) and 50 mM *p*NP α -L-Fuc (concentration in

reaction mixture, dissolved in DMF) were incubated with protein (0.06 mg/mL in reaction mixture) at 45 °C for 3 h. 5% propanol solution (concentration in reaction mixture) and 50 mM *p*NP α -L-Fuc (concentration in reaction mixture) dissolved in DMF were incubated with protein (0.06 mg/mL in reaction mixture) at 45 °C for 3 h. Ethanol, 2-propanol, butanol, pentanol, and octanol were tested as follows: 10% alcohol solution (concentration in reaction mixture) and 50 mM *p*NP α -L-Fuc (concentration in reaction mixture) dissolved in DMF were incubated with protein (0.06 mg/mL in reaction mixture) at 45 °C for 3 h. Conditions for testing *p*-nitrophenylated saccharides: 33 mM *p*NP α -D-Man, *p*NP α -D-Glc, or *p*NP α -D-Gal (concentration in reaction mixture, dissolved in DMF) and 50 mM *p*NP α -L-Fuc (concentration in reaction mixture) dissolved in DMF, were incubated with protein (0.06 mg/mL in reaction mixture) at 45 °C for 3 h. Testing of *p*NP α -L-Fuc functioning simultaneously as an acceptor and donor molecule: 80 mM *p*NP α -L-Fuc (concentration in reaction mixture) dissolved in DMF, was incubated with protein (0.06 mg/mL in reaction mixture) at 45 °C for 3 h.

Thin-layer chromatography

Substrate specificity and transglycosylation assays were analysed using TLC. 1 μ L of each reaction mixture was spotted on silica gel aluminium plates (Sigma Aldrich, USA) alongside with L-fucose, *p*NP α -L-Fuc and L-fucose acceptor compounds. For substrate specificity assay, the reaction mixture was applied on the TLC plate immediately after the start of the reaction (as a control) and after 3 h. L-Fucose was used as a marker. Plates containing saccharides as acceptors and plates from substrate specificity assay were developed in ethyl acetate/acetic acid/water (6:1:1). Plates with the rest of acceptor compounds were developed in butanol/ethanol/water (7:2:2). Chromatograms were stained using 0.1 M 2-methylresorcinol in 5% (v/v) solution of sulphuric acid in ethanol and visualized by heating.

Mass spectrometry analysis

Products of transglycosylation produced in kinetic assays using 2FL as substrate (concentration of 2FL was 25 mM, and 250 mM) were analysed using mass spectrometry analysis in mode ESI+ (Q-Tof micro mass spectrometer, Waters Micromass, Milford, MA) with the direct inlet. The presence of glucose in the maltose sample was confirmed by direct infusion of the sample to the 15T solarix XR FT-ICR mass spectrometer (Bruker Daltonics, USA) operated in positive mode. The instrument was externally calibrated using clusters of sodium trifluoroacetic acid. Raw data were processed by the DataAnalysis 4.1 software (Bruker Daltonics, USA). Ions at m/z 203.0526 and m/z 219.0265 corresponding to sodium and potassium adducts of glucose were observed. The presence of monosaccharides in the protein sample prior to crystallization was examined by the same approach.

Differential scanning fluorimetry

The melting temperature of α -L-f1wt and α -L-f1W392A was tested using DSF. 4 μ L of protein sample (0.8 mg/mL) was loaded into capillaries (Prometheus NT.48 Series nanoDSF Grade standard Capillaries). The protein sample contained 50 mM phosphate buffer, pH 7.5. The measurement was executed using a Prometheus NT.48 (NanoTemper Technologies GmbH, Germany). The experiment parameters were set as follows: start temperature 25 °C; end temperature 90 °C; excitation power 10% and temperature slope 2.0 °C/min.

Acknowledgments

This publication was supported by the project Czech Infrastructure for Integrative Structural Biology for Human Health [CZ.02.1.01/0.0/0.0/16_013/0001776] from the ERDF; the project BIOCEV [CZ.1.05/1.1.00/02.0109] from the ERDF; institutional support [RVO 86652036]; the Ministry of Education, Youth and Sports of the Czech Republic, Centre of Molecular

Structure, Biocev [LM2015043]. Financial support from specific university research [MSMT No 21-SVV/2018] is acknowledged.

Abbreviations

α -L-f1wt, α -L-fucosidase isoenzyme 1 from *P. thiaminolyticus*; α -L-f1W392A, mutant of α -L-fucosidase isoenzyme 1 from *P. thiaminolyticus*; BIS-TRIS, 2,2-bis(hydroxymethyl)-2,2',2''-nitrilotriethanol; Boc-L-Ser-Ome, *N*-(*tert*-butoxycarbonyl)-L-serine; Boc-L-Thr-Ome, *N*-(*tert*-butoxycarbonyl)-L-threonine; CBM, carbohydrate binding module; DLS, dynamic light scattering; DSF, differential scanning fluorimetry; E, enzyme; EPPS, *N*-(2-hydroxyethyl)piperazine-*N'*-(3-propanesulfonic acid); F, L-fucose; 2FL, 2'-fucosyllactose; GH, glycoside hydrolases, HMO, human milk oligosaccharides; L, lactose; LB, Luria-Bertani; Ni-NTA, nickel-nitrilotriacetic acid; *p*NP α -L-Fuc, *p*-nitrophenyl α -L-fucopyranoside; *p*NP α -D-Gal, *p*-nitrophenyl α -D-galactopyranoside; *p*NP α -D-Glc *p*-nitrophenyl α -D-glucopyranoside; *p*NP α -D-Man, *p*-nitrophenyl α -D-mannopyranoside; PCR, polymerase chain reaction; PEG polyethylene glycol; R.m.s.d., root mean square deviation; RPM, revolutions per minute; SDS PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; TLC, thin layer chromatography; TMaFuc, *T. maritima* α -L-fucosidase.

References

- Adams PD, Afonine, PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, and Zwart PH. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst D*. 66:213–221.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol*. 215:403–410.
- Ashida H, Miyake A, Kiyohara M, Wada J, Yoshida E, Kumagai H, Katayama T, Yamamoto K. 2009. Two distinct α -L-fucosidases from *Bifidobacterium bifidum* are essential for the utilization of fucosylated milk oligosaccharides and glycoconjugates. *Glycobiology*. 19(9):1010–1017.
- Avila JL, Convit J. 1974. Studies on human polymorphonuclear leukocyte enzymes. IV. Intracellular distribution and properties of alpha-L-fucosidase. *J Biochim Biophys Acta*. 358:308–318.
- Becker DJ, Lowe JB. 2003. Fucose: biosynthesis and biological function in mammals. *Glycobiology*. 13(7): 41R–53R.

- Benešová E, Lipovová P, Dvořáková H, Králová B. 2013. α -L-Fucosidase from *Paenibacillus thiaminolyticus*: Its hydrolytic and transglycosylation abilities. *Glycobiology*. 23(9):1052–1065.
- Birnboim HC, Doly J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7(6):1513–23.
- Bode L. 2012. Human milk oligosaccharides: Every baby needs a sugar mama. *Glycobiology*. 22(9):1147–1162.
- Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J.* 382:769–781.
- Cao H, Walton JD, Brumm P, Philips GN, Jr. 2014. Structure and substrate specificity of a Eukaryotic fucosidase from *Fusarium graminearum*. *J Biol Chem.* 289(37):25624–25638.
- Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Cryst D.* 66:12–21.
- Cobucci-Ponzano B, Conte F, Rossi M, Moracci M. 2008. The α -L-fucosidase from *Sulfolobus solfataricus*. *Extremophiles*. 12:61–68.
- Davies GJ, Planas A, Rovira C. 2012. Conformational analyses of the reaction coordinate of glycosidases. *Acc Chem Res.* 45(2):308–316.
- Dische Z, di Sant'Agnese P, Pallavicini C, Youlos J. 1959. Composition of mucoprotein fractions from duodenal fluid of patients with cystic fibrosis of the pancreas and from controls. *Pediatrics*. 24:74–91.

- Emsley P, Lohkamp B, Scott W, Cowtan K. 2010. Features and development of Coot. *Acta Cryst D*. 66:486–501.
- Evans PR, Murshudov GN. 2013. How good are my data and what is the resolution? *Acta Cryst D*. 69:1204–1214.
- Guillotin L, Lafite P, Daniellou R. 2014. Unraveling the Substrate Recognition Mechanism and Specificity of the Unusual Glycosyl Hydrolase Family 29 BT2492 from *Bacteroides thetaiotaomicron*. *Biochemistry*. 53:1447–1455.
- Gornik I, Maravić G, Dumić J, Flögel M, Lauc G. 1999. Fucosylation of IgG heavy chains is increased in rheumatoid arthritis. *Clin Biochem*. 32(8):605–608.
- Henrissat B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*. 280:309–316.
- Jabs A, Weiss MS, Hilgenfeld R. 1999. Non-proline *Cis* Peptide Bonds in Proteins. *J Mol Biol*. 286:291–304.
- Kabsch W. 2010. Integration, scaling, space-group assignment and post-refinement. *Acta Cryst D*. 66:133–144.
- Kawai R, Igarashi K, Kitaoka M, Ishii T, Samejima M. 2004. Kinetics of substrate transglycosylation by glycoside hydrolase family 3 glucan (1→3)- β -glucosidase from the white-rot fungus *Phanerochaete chrysosporium*. *Carbohydrate Research*. 339:2851–2857.
- Kirschner KN, Yongye AB, Tschampel SM, Daniels CR, Foley BL, Woods RJ. 2008. GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J Comput Chem*. 29: 622–655.

- Korb O, Stützle T, Exner TE. 2009. Empirical Scoring Functions for Advanced Protein-Ligand Docking with PLANTS. *J Chem Inf Model.* 49:84–96.
- Koshland DE. 1953. Stereochemistry and the mechanism of enzymatic reactions. *Biol Rev Camb Philos Soc.* 28:416–436.
- Krissinel E, Henrick K. 2007. Inference of macromolecular assemblies from crystalline state. *J Mol Biol.* 372:774–797.
- Krissinel E, Henrick K. 2004. Secondary structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Cryst D.* 60:2256–2268.
- Lammerts van Bueren A, Ardèvol A, Fayers-Kerr J, Luo B, Zhang Y, Sollogoub M, Blériot Y, Rovira C, Davies GJ. 2010. Analysis of the Reaction Coordinate of α -L-Fucosidases: A Combined Structural and Quantum Mechanical Approach. *J Am Chem Soc.* 132:1804–1806.
- Liu TW, Ho CW, Huang HH, Chang SM, Popat SD, Wang YT, Wu MS, Chen YJ, Lin CH. 2009. Role for α -L-fucosidase in the control of *Helicobacter pylori*-infected gastric cancer cells. *PANS.* 106(34):14581–14586.
- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The Carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42:D490–D495.
- Lovering AL, Lee SS, Kim YW, Withers SG, Strynadka NCJ. 2005. Mechanistic and structural analysis of a family 31 α -glycosidase and its glycosyl-enzyme intermediate. *J Biol Chem.* 280(3):2105–2115.
- Ma B, Simala-Grant JL, Taylor DE. 2006. Fucosylation in prokaryotes and eukaryotes. *Glycobiology.* 16(12):158R–184R.

McCarter JD, Withers SG. 1994. Mechanisms of enzymatic glycoside hydrolysis. *Curr Opin Struct Biol.* 4:85–92.

Mayes HB, Broadbelt LJ, Backham GT. 2014. How Sugars Pucker: Electronic Structure Calculations Map the Kinetic Landscape of Five Biologically Paramount Monosaccharides and Their Implications for Enzymatic Catalysis. *J Am Chem Soc.* 136:1008–1022.

Michalski JC, Klein A. 1999. Glycoprotein lysosomal storage disorders: α - and β -mannosidosis, fucosidosis and α -N-acetylgalactosaminidase deficiency. *Biochim Biophys Acta.* 1455:69–84.

Miyoshi E, Moriwaki K, Terao N, Tan CC, Terao M, Nakagawa T, Matsumoto H, Shinzaki S, Kamada Y. 2012. Fucosylation Is a Promising Target for Cancer Diagnosis and Therapy. *Biomolecules.* 2:34–45.

Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA. 2011. REFMAC5 for refinement of macromolecular crystal structure. *Acta Cryst D.* 67:355–367.

Nguyen DHD, Park JT, Shim JH, Tran PL, Oktavina EF, Nguyen TLH, Lee SJ, Park CS, Li D, Park SH, Stapleton D, Lee JS, Park KH. 2014. Reaction Kinetics of Substrate Transglycosylation Catalyzed by TreX of *Sulfolobus solfataricus* and Effects on Glycogen Breakdown. *Journal of Bacteriology.* 196:1941–1949.

Reglero A, Cabezas JA. 1976. Glycosidases of Molluscs. *Eur J Biochem.* 66:379–387.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* 25(13):1605–1612.

Pickard JM, Maurice CF, Kinnebrew MA, Abt MC, Schenten D, Golovkina TV, Bogatyrev SR, Ismagilov RF, Pamer EG, Turnbaugh PJ, Chervonsky AV. 2014. Rapid fucosylation of intestinal epithelium sustains host-commensal symbiosis in sickness. *Nature*. 514:638–641.

Robert X, Gouet, P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucl Acids Res*. 42(W1): W320–W324.

Sakurama H, Fushinobu S, Hidaka M, Yoshida E, Honda Y, Ashida H, Kitaoka M, Kumagai H, Yamamoto K, Katayama T. 2012. 1,3-1,4- α -L-Fucosynthase That Specifically Introduces Lewis a/x Antigens into Type -1/2 Chains. *J Biol Chem*. 287(20):16709–16719.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. 7:539.

Skálová T, Dohnálek J, Spiwok V, Lipovová P, Vondráčková E, Petroková H, Dušková J, Strnad H, Králová B, Hašek J. 2005. Cold-active β -Galactosidase from *Arthrobacter* sp. C2-2 Forms Compact 660 kDa Hexamers: Crystal Structure at 1.9 Å Resolution. *J Mol Biol*. 353: 282–294.

Staudacher E, Altmann F, Wilson IB, März L. 1999. Fucose in *N*-glycans: from plant to man. *Biochim Biophys Acta*. 1473:216–236.

Sulzenbacher G, Bignon C, Nishimura T, Tarling CA, Withers SG, Henrissat B, Bourne Y. 2004. Crystal Structure of *Thermotoga maritima* α -L-Fucosidase. Insight into the catalytic mechanism and the molecular basis for fucosidosis. *J Biol Chem*. 279:13119–13128.

Summers EL, Moon CD, Atua R, Arcus VR. 2016. The structure of a glycoside hydrolase 29 family member from a rumen bacterium reveals unique, dual carbohydrate-binding domain. *Acta Cryst F*. 72:750–761.

Vagin A, Teplyakov A. 2010. Molecular replacement with MOLREP. *Acta Cryst D*. 66:22–25.

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. 2011. Overview of the CCP4 suite and current developments. *Acta Cryst D*. 67:235–242.

Wright DW, Moreno-Vargas AJ, Carmona AT, Robina I, Davies GJ. 2013. Three dimensional structure of a bacterial α -L-fucosidase with a 5-membered iminocyclitol inhibitor. *Bioorg Med Chem*. 21:4751–4754.

Yang Y, Qian J, Ming D. 2015. Docking polysaccharides to proteins that have a Tryptophan box in the binding pocket. *Carbohydr Res*. 414:78–84.

Zhang SY, Lin BD, Li BR. 2015. Evaluation of the diagnostic value of alpha-L-fucosidase, alpha-fetoprotein and thymidine kinase 1 with ROC and logistic regression for hepatocellular carcinoma. *FEBS Open Bio*. 5:240–244.

Legends to figures

Fig. 1: Amino acid sequence and secondary structure elements of α -L-f1wt. α -Helices and β -strands are labelled and numbered (α ; β). 3_{10} helices are marked as η and β -turns as T. Calculated based on chain A using PDBFold (Krissinel and Henrick 2004). Catalytic acid/base and nucleophile amino acid residues of the active site are marked with a star, tryptophan complementing the active site of the neighbouring chain is marked with a black circle. The C-terminal domain with the carbohydrate binding module-like fold is marked. The

produced protein construct contained the following sequence at the N-terminus: MGHHHHHHHHHSSGHIEGRH followed by MTLT....

Fig. 2: A) Monomer of α -L-f1wt in secondary structure representation. The N-terminal domain is in blue and the C-terminal domain in pink, the catalytic active site residues Asp186 (nucleophile) and Glu239 (acid/base) are shown in red. The active site residue Trp392 located in the C-terminal domain complementing the active site in the adjacent monomer is shown in grey sticks. The N-termini and C-termini are marked with the corresponding letter. B) Formation of the dimer: the N-terminal domain of chain A is coloured blue, the C-terminal domain pink, chain B is in cyan (N-terminal) and violet (C-terminal). The dimer of α -L-f1wt is formed by contacts of the C-terminal domain of one monomer and the N-terminal domain of the second monomer (view along the non-crystallographic 2-fold axis). C) The hexameric assembly of α -L-f1wt shown from three different views, left panel: along the non-crystallographic 3-fold axis of the hexamer; middle panel: along the non-crystallographic 2-fold axis viewing the enzyme dimer; right panel: along the 2-fold axis rotated by 60°. The individual chains are differentiated by colour and labelled by the corresponding underlined letter (chain label in the PDB entry). The maltose binding sites are illustrated using yellow spheres representation of the maltose molecules and marked by numbers. The active sites are marked in two different views (middle and right panel). Illustrations were created with Pymol (version 1.1r.; Schrödinger, LLC).

Fig. 3: Selected parts of sequence alignment of α -L-f1wt and eight α -L-fucosidases from the family GH29 with structures available arranged by sequence identity to α -L-f1wt. Identical amino acids are in black columns and white letters, conservative changes are in white frames and bold letters. Amino acids of the α -L-f1wt active site are marked with an asterisk. The conserved active site residues are marked with an underlined asterisk. The active site residue conserved only in prokaryotic α -L-fucosidases is marked with an underlined asterisk and an

exclamation mark. The catalytic acid/base residue is marked by an asterisk with letter “a” and in the case of α -L-f1wt placed in a black frame. These residues are all in the same position in the 3D structure, however, they are not identically placed in the enzyme sequence. The residue Trp392 complementing the active site of an adjacent monomer is in a circle and the position in alignment is marked with an asterisk. Sequence numbering is according to α -L-f1wt. Sequences were aligned using Clustal Omega (Sievers et al. 2011) and the illustration was generated using ESPript 3 (Robert and Gouet 2014) and manually edited.

Fig. 4: The active site of α -L-f1wt in the crystal structure. For transparency only selected active site residues are shown in all graphics. Graphics were created with Pymol (version 1.1r.; Schrödinger, LLC). A) Glucose molecule (carbon atoms in black) bound in the active site shown with selected amino acid residues (nucleophile Asp186 and acid/base Glu239 with carbon atoms in dark red; Trp392 complementing the active site from an adjacent monomer is shown with carbon atoms in cyan). B) Surface representation of the α -L-f1wt active site with glucose and selected active site residues in sticks. C) mF_o-DF_c electron density (green) contoured at 2σ prior to introducing the glucose molecule in the model. D) $2F_o-F_c$ composite omit map (black) shown around the glucose molecule contoured at 1σ (PHENIX, Adams et al. 2010).

Fig. 5: Dependence of initial rate on substrate concentration. Reactions catalysed by α -L-f1wt (black circle) and α -L-f1W392A (square), were measured at 37 °C. Error bars represent standard deviations calculated from three replicates of experiment. Data were fitted using the equation for simultaneous hydrolysis and transglycosylation (Nguyen et al. 2014; Kawai et al. 2004) using GraphPad Prism 7.04 for Windows, GraphPad Software, La Jolla California USA. A) Measurement with $pNP\alpha$ -L-Fuc as substrate. B) Measurement with 2FL as substrate.

Fig. 6: Dimer formation and hexameric assembly of α -L-f1wt and TMaFuc. Monomers forming a dimer are coloured with the same colour in a different shade. In α -L-f1wt individual chains are marked by letters according to the PDB entry and underlined using the corresponding colour. TMaFuc chains are marked according to the PDB entry except for chains D-F, which are symmetry-generated mates. A) The hexameric assembly of α -L-f1wt. B) The hexameric assembly of TMaFuc. C) Formation of a dimer within the hexamer of α -L-f1wt. D) Dimer of TMaFuc. In both dimers the N-terminal domain (chain A blue, chain B cyan) and the C-terminal domain (chain A pink, chain B violet) are differentiated by colour. The catalytic residues are represented by spheres in red. Trp392 is represented by grey spheres. The N- and C-termini are coloured yellow and marked. The illustrations were made using Pymol (version 1.1r.; Schrödinger, LLC).

Fig. 7: Superposition of α -L-f1wt (N-terminal domain – blue, C-terminal domain – pink) and TMaFuc (N-terminal domain – orange, C-terminal domain – green). The catalytic active site residues are shown as spheres (α -L-f1wt in red, TMaFuc in light green). Trp392 of α -L-f1wt, complementing the active site in the adjacent monomer, is shown in the C-terminal domain of α -L-f1wt in grey spheres. The N- and C-termini of α -L-f1wt and TMaFuc are in yellow and marked by the corresponding letter underlined in the corresponding colour. The monomers were superimposed using the secondary structure matching algorithm in Coot (Emsley et al., 2010). The Illustration was created with Pymol (version 1.1r.; Schrödinger, LLC).

Fig. 8: Comparison of the active site complementation and the hexameric assembly of α -L-f1wt and α -xylosidase from *E. coli*. A) Left panel: the hexameric assembly of α -L-f1wt, view along the non-crystallographic 3-fold; middle panel: view along the non-crystallographic 2-fold axis, complementing Trp392 is in magenta; right panel: selected residues of the active site; catalytic amino acid residues are shown with carbon atoms in dark red, complementing Trp392 in a magenta circle and with carbon atoms in cyan . B) Left panel: the hexameric

assembly of *E. coli* α -xylosidase, view along the non-crystallographic 3-fold; middle panel: view along the non-crystallographic 2-fold, complementing Trp8 is in magenta; right panel: selected active site residues showing a similar complementation with Trp8 (carbon atoms in cyan, sticks encircled in magenta) from chain B; additional residues from chain C are also complementing the active site (Leu48 and Asp49, carbon atoms in dark green), the catalytic amino acid residues are shown with carbon atoms in dark red. Illustrations created with Pymol (version 1.1r.; Schrödinger, LLC) and manually edited.

Fig. 9: A) Rate equation used for calculating kinetic parameters for hydrolysis and substrate transglycosylation (V_{lim1} and K_M are hydrolysis parameters; V_{lim2} and K_{M2} are transglycosylation parameters). B) The minimal kinetic scheme for simultaneously occurring hydrolysis and transglycosylation of double-displacement catalytic mechanism (E enzyme, 2FL 2'-fucosyllactose, L lactose, F L-fucose).

Tables

Table I. Summary of the known structures of α -L-fucosidases from family GH29.









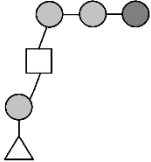
Organism	Protein	PDB ID (reference)	Sequence Identity to α -L-f1wt (%)	Number of domains		Oligomeric state
				TIM barrel	β -sandwich	
<i>P. thiaminolyticus</i>	α -L-f1wt	6GN6	100	1 ^d	1 ^e	hexamer ^b
<i>B. thetaiotaomicron</i>	BT 2970 ^a	2WVS (Lammerts van Bueren et al. 2010)	33	1 ^d	1 ^e	monomer ^c
<i>B. thetaiotaomicron</i>	BT 2192 ^a	4OZO (Guillotin et al. 2014)	28	1 ^d	1 ^e	dimer ^b
<i>B. thetaiotaomicron</i>	BT 3798 ^a	3GZA To Be Published	28	1 ^d	1 ^e	dimer ^c
<i>T. maritima</i>	TMaFuc ^a	1ODU (Sulzenbach er et al. 2004)	29	1 ^d	1 ^e	hexamer ^b
<i>B. longum</i>	Blon_2336 ^a	3UES (Sakurama et al. 2012)	26	1 ^d	1 ^e	dimer ^b
<i>B. ovatus</i>	BACOVA_04357 ^a	4ZRX To Be Published	30	1 ^d	2 ^f	monomer ^c
<i>F. graminearum</i>	FgFCO1 ^a	4PSP (Cao et al. 2014)	23	1 ^d	2 ^f	monomer ^b
<i>unknown bacterium</i>	GH29_0940 ^a	5K9H (Summers et al. 2016)	32	1 ^d	2 ^f	monomer ^c

^a According to CAZy (Lombard et al. 2014). ^b Experimentally confirmed oligomeric state in solution. ^c Oligomeric state according to the PDB entry. ^d N-terminal domain. ^e C-terminal domain. ^f One C-terminal domain and one central domain. Protein sequence identity calculated by BLAST (Altschul et al. 1990).

Table II. Kinetic parameters of α -L-f1wt and α -L-f1W392A for p NP α -L-Fuc and 2FL as substrates calculated from the data in Figure 5 using the equation in Figure 9.

α -L-f1wt		
Substrate	K_M (mmol/L)	V_{lim} (μ mol/min/mg)
p NP α -L-Fuc	0.6 ± 0.1	346 ± 6
2FL	20 ± 4	24 ± 2
α -L-f1W392A		
Substrate	K_M (mmol/L)	V_{lim} (μ mol/min/mg)
p NP α -L-Fuc	5.8 ± 0.5	320 ± 20
2FL	19 ± 2	7.0 ± 0.3

Table III. The summary of the used substrates and results of substrate specificity experiments.

Substrate	Structure	Type of bond	α -L-f1wt	α -L-f1W392A
2'-Fucosyllactose (Fuc- α -1,2-Gal- β -1,4-Glc)		α -1,2	++++	++
3-Fucosyllactose (Gal- β -4(Fuc- α -1,3)Glc)		α -1,3	++	+
Lewis X trisaccharide (Gal- β -1,4(Fuc- α -1,3) GlcNAc)		α -1,3	-	-
Lewis A trisaccharide methyl glycoside (Gal- β -1,3(Fuc- α -1,4) GlcNAc- β -OMe)		α -1,4	+	a+
Blood group H disaccharide (Fuc- α -1,2-Gal)		α 1,2	++++	++
2-Acetamido-2-deoxy-3-O- (α -L-fucopyranosyl)-D- glucopyranose (Fuc- α -1,3-GlcNAc)		α -1,3	++++	++
2-Acetamido-2-deoxy-4-O- (α -L-fucopyranosyl)-D- glucopyranose (Fuc- α -1,4-GlcNAc)		α -1,4	+++	++
2-Acetamido-2-deoxy-6-O- (α -L-fucopyranosyl)-D- glucopyranose (Fuc- α -1,6-GlcNAc)		α -1,6	+++	++
Globo H (Fuc- α -1,2-Gal- β -1,3- GalNAc- β -1,3-Gal- α -1,4- Gal- β -1,4-Glc)		α -1,2	+	a+

■ *N*-acetylglucosamine, \triangle L-fucose, \circ galactose, \bullet glucose, \square *N*-acetylgalactosamine, \blacksquare

methyl *N*-acetylglucosamine, the number of characters “+” indicates relative activity, + the

substrate was hydrolysed, - the substrate was not hydrolysed, a+ hydrolysed about a half of the amount of the substrate when compared to +.

Table IV. Summary of transglycosylation reaction results for α -L-f1W392A, monitored by TLC.

Acceptor	Transglycosylation	Acceptor	Transglycosylation
<i>p</i> NP α -L-Fuc	+	1-Octanol	+
<i>p</i> NP α -D-Gal	+	Boc-L-Ser-OMe	+
<i>p</i> NP α -D-Glc	+	Boc-L-Thr-OMe	+
<i>p</i> NP α -D-Man	+	D-Glucose	+
Methanol	+	D-Galactose	+
Ethanol	+	D-Mannose	-
1-Propanol	+	D-Fructose	-
2-Propanol	+	D-Glucosamine	-
Butanol	+	<i>N</i> -acetyl-D-Glucosamine	-
Pentanol	+	D-Maltose	+

+ Transfer of L-fucosyl moiety successful, - transfer of L-fucosyl moiety unsuccessful, in all cases the donor of the L-fucosyl moiety was *p*NP α -L-Fuc.

Table V. Data collection statistics and refinement parameters for α -L-f1wt. Statistics in parentheses refer to highest resolution shell.

Data collection statistics for α-L-f1wt	
PDB ID	6GN6
Wavelength (Å)	0.9201
Crystal-to-detector distance (mm)	419.82
No. of oscillation images processed	1330
Oscillation width (°)	0.1
Exposure time per image (s)	0.1
Space group	$P2_12_12_1$
Unit cell parameters (Å)	$a=107.77$ $b=148.76$ $c=194.90$
Resolution (Å)	48.93-2.20 (2.24-2.20)
No. of observations	783892 (36540)
No. of unique reflections	157780 (7722)
Average mosaicity (°)	0.16
Multiplicity ($I+=I-$)	5.0 (4.7)
R_{merge}	0.099 (0.709)
R_{pim}	0.050 (0.351)
R_{meas}	0.112 (0.795)
Mean I/σ_I	12.3 (2.1)
Completeness (%)	99.4 (99.2)
CC1/2	0.996 (0.724)
Wilson B (Å ²)	34.2
Refinement parameters	
R_{work}	0.1673
R_{free}	0.2173
R.m.s.d. bonds (Å)	0.016
R.m.s.d. angles (°)	1.825
Ramachandran plot (%) ^a	
Favored (%)	96
Outliers (%)	0
Average B (Å ²)	

All atoms	35.3
Main chain	31.6
Side chain	38.3
Ligands	65.7
Solvent	38.3

^a Calculated by MolProbity (Chen et al. 2010)