



## Substrate specificity and transufcosylation activity of GH29 $\alpha$ -l-fucosidases for enzymatic production of human milk oligosaccharides

**Zeuner, Birgitte; Muschiol, Jan; Holck, Jesper; Lezyk, Mateusz; Gedde, Mattias Raae; Jers, Carsten; Mikkelsen, Jørn Dalgaard; Meyer, Anne S.**

*Published in:*  
New Biotechnology

*Link to article, DOI:*  
[10.1016/j.nbt.2017.12.002](https://doi.org/10.1016/j.nbt.2017.12.002)

*Publication date:*  
2018

*Document Version*  
Peer reviewed version

[Link back to DTU Orbit](#)

### *Citation (APA):*

Zeuner, B., Muschiol, J., Holck, J., Lezyk, M., Gedde, M. R., Jers, C., Mikkelsen, J. D., & Meyer, A. S. (2018). Substrate specificity and transufcosylation activity of GH29  $\alpha$ -l-fucosidases for enzymatic production of human milk oligosaccharides. *New Biotechnology*, 41, 34-45. <https://doi.org/10.1016/j.nbt.2017.12.002>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

## Accepted Manuscript

Title: Substrate specificity and transglycosylation activity of GH29  $\alpha$ -L-fucosidases for enzymatic production of human milk oligosaccharides

Authors: Birgitte Zeuner, Jan Muschiol, Jesper Holck, Mateusz Lezyk, Mattias Raae Gedde, Carsten Jers, Jørn Dalgaard Mikkelsen, Anne S. Meyer



PII: S1871-6784(17)30460-0  
DOI: <https://doi.org/10.1016/j.nbt.2017.12.002>  
Reference: NBT 1038

To appear in:

Received date: 24-8-2017  
Revised date: 29-11-2017  
Accepted date: 4-12-2017

Please cite this article as: Zeuner, Birgitte, Muschiol, Jan, Holck, Jesper, Lezyk, Mateusz, Gedde, Mattias Raae, Jers, Carsten, Mikkelsen, Jørn Dalgaard, Meyer, Anne S., Substrate specificity and transglycosylation activity of GH29  $\alpha$ -L-fucosidases for enzymatic production of human milk oligosaccharides. *New Biotechnology* <https://doi.org/10.1016/j.nbt.2017.12.002>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Substrate specificity and transfuosylation activity of GH29  $\alpha$ -L-fucosidases for enzymatic production of human milk oligosaccharides**

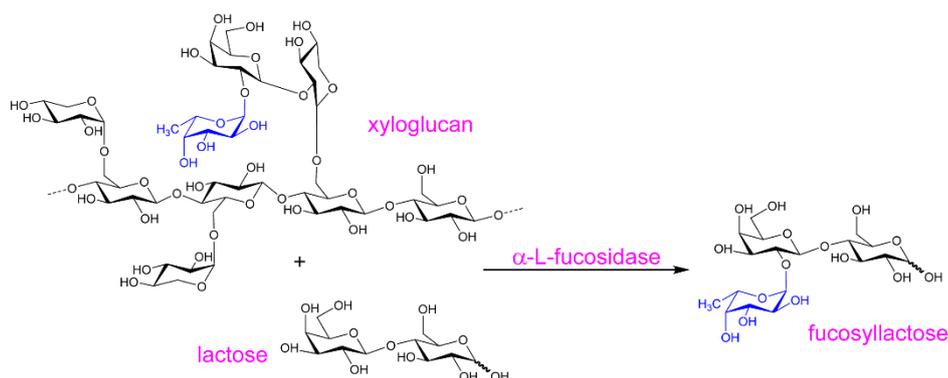
Birgitte Zeuner\* biz@kt.dtu.dk, Jan Muschiol, Jesper Holck, Mateusz Lezyk, Mattias Raae Gedde, Carsten Jers, Jørn Dalgaard Mikkelsen, Anne S. Meyer

Center for Bioprocess Engineering, Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 229, DK-2800 Kgs. Lyngby

\*Corresponding author:

ACCEPTED MANUSCRIPT

## Graphical Abstract

**Highlights:**

- Seven GH29  $\alpha$ -L-fucosidases all possess transglycosylation activity
- $\alpha$ -L-Fucosidase from *F. graminearum* catalyses formation of 2'FL from xyloglucan
- $\alpha$ -L-Fucosidase CpAfc2 from *C. perfringens* efficiently catalyses formation of LNFP II
- Systematic characterization of substrate specificity for seven GH29  $\alpha$ -L-fucosidases

**Abstract**

Human milk oligosaccharides (HMOs) constitute a unique family of bioactive lactose-based molecules present in human breast milk. HMOs are of major importance for infant health and development but also virtually absent from bovine milk used for infant formula. Among the HMOs, the fucosylated species are the most abundant. Transglycosylation catalysed by retaining  $\alpha$ -L-fucosidases is a new route for manufacturing biomimetic HMOs. Seven  $\alpha$ -L-fucosidases from glycosyl hydrolase family 29 were expressed, characterized in terms of substrate specificity and thermal stability, and shown to be able to catalyse transglycosylation. The  $\alpha$ -L-1,3/4-fucosidase CpAfc2 from *Clostridium perfringens* efficiently catalysed the formation of the more complex human milk oligosaccharide structure lacto-*N*-fucopentaose II (LNFP II) using 3-fucosyllactose as fucosyl donor and lacto-*N*-tetraose as acceptor with a 39% yield.  $\alpha$ -L-Fucosidases FgFCO1 from *Fusarium graminearum* and Mfuc5 from a soil metagenome were able to catalyse transglycosylation of lactose using citrus xyloglucan as fucosyl donor. FgFCO1 catalysed formation of 2'-fucosyllactose, whereas Mfuc5 catalysis mainly produced an unidentified, non-HMO fucosyllactose, reaching molar yields based on the donor substrate of 14% and 18%, respectively.

**Abbreviations:** 2'FL: 2'-fucosyllactose; 3FL: 3-fucosyllactose; A:D: acceptor-to-donor ratio; BbAfcB:  $\alpha$ -1,3/4-L-fucosidase from *Bifidobacterium bifidum*; CpAfc2:  $\alpha$ -1,3/4-L-fucosidase from *Clostridium perfringens*; FgFCO1:  $\alpha$ -L-fucosidase from *Fusarium graminearum*; FL: fucosyllactose; Fuc: fucose; GH: glycosyl hydrolase; GlcNAc: *N*-acetylglucosamine; HMO: human milk oligosaccharide; IMAC: immobilized metal affinity chromatography; LacNAc: *N*-acetyllactosamine; LNT: lacto-*N*-tetraose; LNnT: lacto-*N*-neotetraose;

LNFP: lacto-*N*-fucopentaose; Mfuc5:  $\alpha$ -L-fucosidase from a soil metagenome; NixE:  $\alpha$ -L-fucosidase from *Xanthomonas campestris* pv. *campestris*; pNP- $\alpha$ -L-Fuc: *p*-nitrophenyl  $\alpha$ -L-fucopyranoside; TfFuc1:  $\alpha$ -1,2-L-fucosidase from *Tannerella forsythia*; TmaFuc:  $\alpha$ -L-fucosidase from *Thermotoga maritima*; XXFG: defined nonasaccharide from xyloglucan; XXLG: defined octasaccharide from xyloglucan (defucosylated XXFG); XyG: xyloglucan.

*Keywords:*  $\alpha$ -L-Fucosidase; GH29; human milk oligosaccharides; substrate specificity; transfucosylation; xyloglucan **Introduction**

Human milk oligosaccharides (HMOs) constitute a unique family of bioactive, heterogeneous carbohydrate oligomers present in human breast milk (5-15 g/L), which are known to be crucial for infant health and development through immunomodulatory effects as well as through selective stimulation of certain *Bifidobacterium* strains [1, 2, 3]. In brief, almost all HMO molecules have lactose (Gal- $\beta$ 1,4-Glc) at the reducing end and are biologically synthesized via sequential elongations of *N*-acetylglucosamine (GlcNAc) and galactose (Gal) moieties. The HMO backbone is then further modified with sialic acid (*N*-acetylneuraminic acid) and/or fucose (Fuc) substitutions [1]. Over 180 different HMO species with different backbone lengths, types of substitutions, and bond configurations are known to exist – the vast majority being fucosylated species [4-6]. Only very few HMO species are present in bovine milk and only at infinitesimal concentrations [1], which is a challenge in production of valuable infant formula. Instead, prebiotic galactooligosaccharides (GOS) and fructooligosaccharides (FOS) are added to infant formula, but these are structurally different from the oligosaccharides occurring in milk and may therefore not elicit the same immune responses [2, 7]. Hence, a large incentive exists for biosynthetic production of authentic HMO structures for use in infant formula as well as in functional studies.

While metabolic engineering of *Escherichia coli* has recently been shown to allow direct microbial production of appreciable fucosyllactose (FL) yields [8-10], production of structurally more complex HMOs in engineered *E. coli* is still in its infancy [10, 11]. As a simpler alternative, direct enzymatic regiospecific synthesis of fucosylated HMOs *in vitro* employing  $\alpha$ -L-fucosidases to catalyze transfucosylation is an option [12-15]. Retaining  $\alpha$ -L-fucosidases are classified in glycosyl hydrolase (GH) family 29 (GH29), whereas inverting  $\alpha$ -L-fucosidases are classified in GH95 (www.cazy.org [16]). For transglycosylation, only retaining glycosidases are relevant [17]. GH29 has been divided into two subfamilies – family GH29A and family GH29B – based on sequence homology which in turn designates substrate specificity. Family GH29A contains  $\alpha$ -L-fucosidases with a relatively relaxed fucosyl substrate specificity, whereas the  $\alpha$ -L-fucosidases in family GH29B are more regiospecific, targeting  $\alpha$ -1,3- and  $\alpha$ -1,4-linkages with a branched Gal only [18]. Several specific amino acids in the active site appear to be responsible for this difference in substrate

specificity [18, 19]. Consequently, regioselectivity in the transfucosylation reaction is more easily controlled with members of the GH29B subfamily [15, 19], but the strict regiospecificity also hinders the production of  $\alpha$ -1,2-fucosylated glycans via transfucosylation catalysed by GH29B enzymes.

So far,  $\alpha$ -L-fucosidase-catalysed production of HMOs has been limited mainly by the lack of appropriate, fucosylated donor substrates [12, 14]. One natural suitable fucosyl donor substrate is plant cell wall derived xyloglucan (XyG), which features terminal  $\alpha$ -1,2-fucosylation on certain side chains, the degree of fucosylation being dependent on the plant source [20, 21]. Only few  $\alpha$ -L-fucosidases have been reported to release Fuc from XyG. Of these, most are inverting GH95  $\alpha$ -L-fucosidases from plants [22, 23], whereas reports on GH29  $\alpha$ -L-fucosidases include only the  $\alpha$ -L-fucosidase FgFCO1 from *Fusarium graminearum* [24, 25] and the  $\alpha$ -L-fucosidase Mfuc5 from a soil metagenome [14]. The  $\alpha$ -L-fucosidase Tm $\alpha$ Fuc from *Thermotoga maritima* has also been hypothesized to be able to release Fuc from XyG [26]; however, when tested alongside Mfuc5 the activity of Tm $\alpha$ Fuc on XyG was low [14]. The present work was undertaken to assess the option for employing microbial GH29 type A and B  $\alpha$ -L-fucosidases for HMO production by transfucosylation using abundantly available industrial food-processing material as substrates, namely XyG extracted from citrus peel and lactose from milk whey. Finally, attempts to establish learning of  $\alpha$ -L-fucosidase specificity were made through studies of enzyme 3D structures and docking of an acceptor substrate.

## Materials & Methods

### Chemicals

The synthetic substrate *p*-nitrophenyl  $\alpha$ -L-fucopyranoside (*p*NP- $\alpha$ -L-Fuc) was purchased from Carbosynth Ltd. (Compton, UK). The fucosylated oligosaccharides 2'-fucosyllactose (2'FL), 3-fucosyllactose (3FL), lacto-*N*-tetraose (LNT), lacto-*N*-neotetraose (LNnT), lacto-*N*-fucopentaose V (LNFP V), and the XyG nonasaccharide XXFG extracted from apple were purchased from Elicityl Oligotech (Crolles, France). L-Fucose,  $\beta$ -lactose, and all other chemicals were purchased from Sigma-Aldrich (Steinheim, Germany). Dried citrus peel was supplied by Dupont Nutrition BioSciences ApS (Brabrand, Denmark) and stored at -20°C until use.

### Preparation of XyG from citrus peel

20 g milled citrus peel was suspended in 1 L of 24% (w/v) KOH with 0.1 % (w/v) NaBH<sub>4</sub> (the latter added to minimize backbone degradation by alkaline 'peeling') and stirred for 24 h [21, 27]. After centrifugation at 5000 *g* for 30 min, the supernatant was neutralized with glacial acetic acid. The suspension was filtered on a 0.22  $\mu$ m filter to remove large, suspended particles. Finally, the solution was diafiltered on a 5 kDa cellulose membrane (Vivaflow 200; Sartorius AG, Göttingen, Germany) against MilliQ water to remove

smaller molecules and salts. The retentate was freeze-dried and subjected to sulfuric acid hydrolysis as described previously [28] and monosaccharide composition was determined by HPAEC-PAD.

### Expression and purification of $\alpha$ -L-fucosidases

Amino acid sequences for FgFCO1, TfFuc1, BbAfcB, CpAfc2, NixE, Mfuc5, and Tm $\alpha$ Fuc (Table 1) were retrieved from GenBank. Signal peptides and transmembrane regions were predicted using the online tools SignalP 4.1 [29] and TMHMM v. 2.0 [30]. Signal peptides were predicted for FgFCO1, TfFuc1, BbAfcB, CpAfc2, and NixE, and were removed from the sequence prior to insertion in the vector (Table 1). A transmembrane region at the C-terminus of BbAfcB was also removed (Table 1).

The resulting genes encoding for FgFCO1, TfFuc1, BbAfcB, CpAfc2, and NixE were codon-optimized for expression in *Escherichia coli* (Table S1), synthesized and inserted into the vector pET22b(+) by GenScript (Piscataway, NJ) using the restriction sites given in Table 1. *E. coli* DH5 $\alpha$  was used for plasmid propagation. *E. coli* BL21 (DE3) or *E. coli* BL21 (DE3) pLysS were transformed with the resulting plasmids and selected for ampicillin resistance (Table 1). For expression, transformants were grown in lysogenic broth (LB) supplemented with 100  $\mu$ g/mL ampicillin at 37°C. When reaching an OD<sub>600</sub> of 0.6-0.9, the temperature was reduced to 25°C (or 18°C for TfFuc1) and expression was induced with 1 mM IPTG. Cells were grown for 20 h before being harvested. Mfuc5 and Tm $\alpha$ Fuc were expressed from the pETM10 vector as described previously [14].

*E. coli* cells were harvested by centrifugation (5300 *g*, 15 min, 5°C) and the pellet was resuspended in binding buffer (20 mM sodium phosphate buffer, 20 mM imidazole, 500 mM NaCl, pH 7.4) and washed by centrifugation (5300 *g*, 15 min, 5°C). Again, the pellet was resuspended in binding buffer followed by sonication (10 rounds of 30 seconds at 100% amplitude and 60% cycle with 30-second breaks in between using a UP400S ultrasonicator (Hielscher Ultrasonics GmbH, Teltow, Germany)) to lyse the cells and centrifugation to remove cell debris (20000 *g*, 20 min, 5°C). The resulting supernatant was passed through a 0.45  $\mu$ m filter before being loaded onto a 4 mL Ni<sup>2+</sup> Sepharose FF column (GE Healthcare, Uppsala, Sweden) for immobilized metal affinity chromatography (IMAC) purification. Unbound material was washed off with 5-7 column volumes (CV) binding buffer. The  $\alpha$ -L-fucosidases were eluted with 3 CV elution buffer (20 mM sodium phosphate buffer, 500 mM imidazole, 500 mM NaCl, pH 7.4) in 0.5 CV fractions. The IMAC-purified  $\alpha$ -L-fucosidase preparations were desalted on PD-10 columns (GE Healthcare, Uppsala, Sweden) using a glycerol-free desalting buffer (20 mM sodium phosphate buffer, 100 mM NaCl, pH 7.4) to avoid interference of glycerol with HPAEC-PAD analysis. Protein purity was confirmed by SDS-PAGE and protein concentration was determined with Bradford reagent (Sigma-Aldrich, Steinheim, Germany) using bovine serum albumin (BSA) standards.

The gene encoding for FgFCO1 with a C-terminal His<sub>6</sub>-tag was codon-optimized for expression in *Pichia pastoris* (Table S1), synthesized and inserted into the vector pPICZ $\alpha$  A by GenScript (Piscataway, NJ) using the restriction sites given in Table 1. After plasmid propagation in *E. coli* DH5 $\alpha$ , the resulting linearized plasmid (prepared according to the EasySelect™ *Pichia* Expression Kit user manual, Invitrogen, 2010) was transformed into *P. pastoris* X-33 by electroporation as described elsewhere [31]. Transformants were selected for Zeocin resistance and grown in buffered glycerol-complex medium (BMGY) for 24 h at 30°C in shake flasks. After centrifugation (2000 *g*, 5 min), the pellet was resuspended in buffered methanol-complex medium (BMMY) and grown for 72 h at 30°C with addition of new methanol (0.5% final concentration) every 24 h. The cultures were harvested by centrifugation (2000 *g*, 5 min), and the supernatant was ultrafiltered on a 10 kDa polyethersulfone membrane (Vivaspin 20; Sartorius AG, Göttingen, Germany). The resulting retentate was mixed 2:1 with concentrated binding buffer (60 mM sodium phosphate buffer, 60 mM imidazole, 1.5 M NaCl, pH 7.4), and the pH was adjusted to 7.4 before IMAC purification, which was carried out on a 5 mL HisTrap HP column (GE Healthcare, Uppsala, Sweden) using an ÄKTA purifier (GE Healthcare, Uppsala, Sweden) with the same binding and elution buffers as above, but using a linear gradient from 0 to 100% elution buffer over 5 CV for elution. The IMAC purified  $\alpha$ -L-fucosidase preparation was desalted and analysed as above. A sample of purified FgFCO1 was treated with Endo H (New England BioLabs Inc., Ipswich, MA) following the manufacturer's instructions and analysed by SDS-PAGE.

### Hydrolytic activity

Hydrolytic activity was determined for all  $\alpha$ -L-fucosidases using a substrate concentration of 0.1 mM fucosylated substrate (*p*NP- $\alpha$ -L-Fuc, 2'FL, 3FL, XXFG, or XyG) in 50 mM buffer (acetate buffer for pH 5.0, phosphate buffer for pH 7.0, and glycine-NaOH buffer for pH 9.0) and enzyme concentrations of 0.05  $\mu$ M or 0.5  $\mu$ M (for Mfuc5 also at 0.005  $\mu$ M). For XyG extracted from citrus peel, the substrate concentration was 0.1 mM bound Fuc as determined by acid hydrolysis (see 2.2). BbAfcB, NixE, FgFCO1, and Tm $\alpha$ Fuc were tested at pH 5.0, whereas CpAfc2, TfFuc1, and Mfuc5 were tested at pH 7.0. TfFuc1 was also tested at pH 9.0 on *p*NP- $\alpha$ -L-Fuc. Negative controls without enzyme were included. All reactions took place at 40°C, were monitored for up to 20 min (180 min for XXFG and XyG), and stopped by mixing a sample 1:1 with 1 M Na<sub>2</sub>CO<sub>3</sub> (resulting pH > 10). For *p*NP- $\alpha$ -L-Fuc, absorbance was measured at 405 nm and calibrated against a *p*-nitrophenol (*p*NP) standard curve. For all other substrates, the amounts of substrate and hydrolysis products were quantified by HPAEC-PAD analysis (see 2.7). Specific hydrolytic activity was defined as  $\mu$ mol product released per minute (U) per  $\mu$ mol of enzyme.

### Thermal stability

Purified  $\alpha$ -L-fucosidases were incubated at temperatures from 40°C to 95°C for relevant time intervals up to 1 h. Afterwards, the remaining hydrolytic activity was measured at standard conditions (40°C, pH 5.0 for NixE, FgFCO1, BbAfcB, and Tm $\alpha$ Fuc, pH 7.0 for Mfuc5, TfFuc1, and CpAfc2) on 0.1 mM substrate (*p*NP- $\alpha$ -L-Fuc for Mfuc5, Tm $\alpha$ Fuc, NixE, and TfFuc1, 3FL for BbAfcB and CpAfc2, and 2'FL for FgFCO1) using enzyme concentrations of 0.05  $\mu$ M (0.5  $\mu$ M for FgFCO1). Reactions were performed as described in 2.4. To ensure linearity in the product release with time, reaction time was 5 min for *p*NP- $\alpha$ -L-Fuc, 10 min for 3FL, and 15 min for 2'FL.

### **Transfucosylation activity**

For transfucosylation of LNT and LNnT, 3FL was used as donor substrate for reactions catalysed by BbAfcB and CpAfc2, whereas 2'FL was used as donor substrate with Tm $\alpha$ Fuc, FgFCO1, Mfuc5, TfFuc1, and NixE (LNT acceptor only). Enzyme concentration was 0.5  $\mu$ M and reaction took place at 40°C for up to 60 min in 100 mM buffer (acetate buffer pH 4.6 for FgFCO1 and Tm $\alpha$ Fuc, pH 5.0 for NixE, pH 5.5 for BbAfcB, phosphate buffer pH 7.0 for CpAfc2, Mfuc5, and TfFuc1). Reactions were terminated by heating at 99°C for 15 min. For reactions with an acceptor-to-donor ratio (A:D) of 10, donor substrate (2'FL or 3FL) concentration was 10 mM and acceptor substrate (LNT or LNnT) concentration was 100 mM. For reactions with an A:D of 1, concentrations of donor and acceptor substrates were 20 mM. Transfucosylation products were identified and quantified by LC-ESI-MS using an external LNFP V standard (see 2.8). To aid the LNFP isomer identification, samples from the reactions were subjected to specific enzymatic hydrolysis as follows: A reaction sample was diluted 10 times in 100 mM acetate buffer (pH 5.0) and subjected to hydrolysis by either BbAfcB or Tm $\alpha$ Fuc at an enzyme concentration of 1  $\mu$ M for 24 h at 40°C. The reaction was terminated by heating at 99°C for 15 min followed by LC-ESI-MS analysis.

For transfucosylation of lactose using the XyG-rich fraction from citrus peel, reactions took place at 40°C for up to 24 h. The concentration of Fuc bound in XyG was 2 mM, the concentration of lactose was 100 mM (A:D = 50), and the enzyme concentration was 2.5  $\mu$ M. All reactions took place in 100 mM acetate buffer (pH 4.6 for FgFCO1 and Tm $\alpha$ Fuc, pH 5.0 for NixE) or 200 mM phosphate buffer (pH 7.0 for Mfuc5 and TfFuc1). Reactions were terminated by heating at 99°C for 15 min. Prior to analysis by HPAEC-PAD (see 2.7), the samples were filtered on a 5 kDa polyethersulfone membrane (Vivaspin 500, Sartorius AG, Göttingen, Germany) to remove large XyG molecules.

### **HPAEC-PAD analysis**

High-performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) was carried out on a Dionex ICS3000 system (Dionex Corp., Sunnyvale, CA) using a CarboPac™ PA1 (4 mm x 250 mm) analytical column equipped with a CarboPac™ PA1 (4 mm x 50 mm) guard column (Dionex Corp.,

Sunnyvale, CA) and a flow rate of 1 mL/min at 25°C. The eluent system comprised MilliQ water (A), 500 mM NaOH (B), and 500 mM NaOAc with 0.02% (w/v) NaN<sub>3</sub> (C).

After enzymatic hydrolysis, 2'FL and 3FL were quantified by isocratic elution at 90:10:0 (% A:B:C) for 30 min using external standards of L-fucose, lactose, 2'FL, and 3FL. Strongly retained anions were washed from the column by elution with 10:10:80 (% A:B:C) for 5 min followed by re-equilibration of the column at 90:10:0 (% A:B:C) for 10 min. Similarly, enzymatic hydrolysis of XyG as well as enzymatic transglucosylation using XyG as donor and lactose as acceptor were quantified by isocratic elution at 95:5:0 (% A:B:C) for 30 min (for XyG hydrolysis) or 45 min (for XyG transglucosylation) using external standards of L-fucose, lactose, 2'FL, and 3FL (3FL was also used for quantification of the unidentified FL formed by Mfuc5). Strongly retained anions were washed from the column by elution with 15:5:80 (% A:B:C) for 5 min followed by re-equilibration of the column at 95:5:0 (% A:B:C) for 10 min. For these analyses, the system was equipped with a post-column setup using 200 mM NaOH at a flow rate of 0.1 mL/min. The retention time of the self-condensation product Fuc-Fuc-*p*NP was determined as outlined previously [14].

Enzymatic hydrolysis of XXFG was quantified by elution with a linear gradient from 80:20:0 (% A:B:C) to 60:20:20 (% A:B:C) from 0 to 30 min using external standards of XXFG and L-fucose. Furthermore, the identity of the resulting XXLG peak was confirmed by LC-MS (see 2.8). The XXLG peak was quantified as XXFG equivalents, and the results were validated against the L-fucose data. Strongly retained anions were removed by elution with 0:20:80 (% A:B:C) for 5 min followed by re-equilibration of the column at 80:20:0 (% A:B:C) for 10 min.

Monosaccharides were quantified using a Dionex ICS-5000 system (DionexCorp, Sunnyvale, CA, USA) equipped with a CarboPac PA1 analytical column (250 × 4 mm) and a CarboPac PA1 guard column (250 × 4 mm) operated at a flow rate of 1 mL/min at 25°C. The eluent system comprised MilliQ water (A) and 500 mM NaOH (B). The elution profile was as follows: isocratic elution 100:0 (% A:B) for 38 minutes, 0:100 (%A:B) for 15 minutes, and re-equilibration 100:0 (% A:B) for 22 minutes. Detection was done by post-column addition of 500 mM NaOH at 0.2 mL/min.

#### **LC-ESI-MS analysis**

Identification and quantification of LNFP transglucosylation products were performed by liquid chromatography electrospray ionization mass spectrometry (LC-ESI-MS) on an Amazon SL iontrap (Bruker Daltonics, Bremen Germany) coupled to an UltiMate 3000 UHPLC from Dionex (Sunnyvale, CA USA). 5 µL sample was injected on a porous graphitized carbon column (Hypercarb PGC, 150 mm × 2.1 mm; 3 µm, Thermo Fisher Scientific, Waltham, MA, USA). The chromatography was performed at 0.4 mL/min at 70°C on a two-eluent system with eluent A (0.1% formic acid in water) and eluent B (acetonitrile). The elution profile was as follows: 0-1 min, 0% B; 1-28 min, linear gradient to 27.5% B; 28-31 min, linear gradient to

60% B; 31-35 min, isocratic 60% B; 35-40 min, isocratic 0% B. The electrospray was operated in negative mode with UltraScan mode and a scan range from 100-1000 m/z, smart parameter setting of 850 m/z, capillary voltage at 4.5 kV, end plate off-set 0.5 kV, nebulizer pressure at 3.0 bar, dry gas flow at 12.0 L/min, and dry gas temperature at 280°C. CID fragmentation of deprotonated ions and formate adducts was performed by Multiple Reaction Monitoring using SmartFrag enhanced amplitude ramping from 80 to 120%, fragmentation time 20 ms. Quantification was performed in Compass QuantAnalysis 2.2 (Bruker Daltonics, Bremen Germany) using LNFP V as external standard. Identification of FL transfucosylation products was carried out in the same manner using 2'FL and 3FL as external standards.

Identification of XXFG-related compounds was performed on the same LC-ESI-MS, with the same eluents, but equipped with a TSK-GEL Amide-80 column (250 x 4.6 mm i.d., 5  $\mu$ m, Tosoh Bioscience, Tokyo, Japan). The chromatography was performed at 1.0 mL/min at 55°C. The samples were diluted to a final concentration of 75% acetonitrile prior to injection. The elution profile was as follows: 0-5 min, 75% B; 5-65 min, linear gradient to 25% B; 65-68 min, isocratic 10% B; 68-75 min, isocratic 75% B. The electrospray was operated in positive mode with a scan range from 100-2000 m/z, smart parameter setting of 1300 m/z, and nebulizer pressure at 5.0 bar. Sodium adducts were detected in auto MS<sup>2</sup>. All other settings as for LNFP compounds.

#### **Homology modelling and substrate docking**

The homology models for all enzymes (except for FgFCO1 and Tm $\alpha$ Fuc) were prepared using the built-in *Homology Modeling* function of YASARA Structure, version 16.9.23 (YASARA Biosciences GmbH, Vienna, Austria) [32]. In that process, the program automatically performs a sequence alignment with the available structure databases and picks the best templates. The covalent fucosyl-enzyme intermediates were prepared by docking  $\alpha$ -L-Fuc into the homology models or available crystal structures respectively (for FgFCO1 (PDB 4NI3) and Tm $\alpha$ Fuc (PDB 2ZWY)). The YASARA dock\_run macro was utilized without changes using a simulation cell of 125  $\text{\AA}^3$  around the active site Asp. The resulting docked models were compared to the TmFuc crystal structure in complex with an inhibitor (PDB 1HL9) to identify the most probable positioning of the ligand. Subsequently, a covalent bond was introduced between C1 of the  $\alpha$ -L-Fuc and the sidechain oxygen of the active site Asp from the opposite side of the  $\alpha$ -OH group using the *Build* function of the program. Afterwards, the  $\alpha$ -OH group was replaced with a hydrogen atom. In this way, the retaining mechanism of GH29 fucosidases is simulated. Finally, the covalent intermediate models were refined using the md\_refine macro of YASARA [33]. The structure with the lowest calculated energy was considered as the final covalent fucosyl-enzyme intermediate model. The resulting structures were analysed for tunnels using the CAVER 3.0 plugin for PyMOL [34]. Docking of acceptor molecules into the covalent intermediate models was carried out as described above except that the simulation cell was positioned around the

fucosyl residue. All figures showing 3D structures were prepared using PyMol (The PyMOL Molecular Graphics System, Version 1.1 Schrödinger, LLC).

### Statistics

One-way ANOVA for determination of statistical significance was performed with JMP®, version 12.1.0 (SAS Institute Inc., Cary, NC). Statistical significance was established at  $p < 0.05$ .

## Results & Discussion

### Expression of $\alpha$ -L-fucosidases

Seven  $\alpha$ -L-fucosidases, which have all been characterized as retaining  $\alpha$ -L-fucosidases in GH29, were selected, expressed, and studied in the current work in order to assess their ability to release Fuc from XyG as well as to investigate their transfucosylation potential (Table 1; Table 2). FgFCO1 and Mfuc5 were included based on their reported activity on XyG [14, 24]. In a phylogenetic tree, the  $\alpha$ -L-fucosidase NixE from *Xanthomonas campestris* pv. *campestris* [35] was adjacent to FgFCO1 [14] and was therefore also included. Having shown activity towards terminal  $\alpha$ -1,2-fucosyl moieties, the  $\alpha$ -1,2-L-fucosidase Tffuc1 from *Tannerella forsythia* [36] also appeared to be a good candidate for XyG utilization. Finally, Tm $\alpha$ Fuc which is probably the most studied  $\alpha$ -L-fucosidase was included as a benchmark enzyme for these GH29A  $\alpha$ -L-fucosidases [26, 37, 38]. In order to also study members of the GH29B subfamily and determine their potential for regioselective production of more complex HMO structures such as LNFP II and LNFP III, the  $\alpha$ -1,3/4-L-fucosidase BbAfcB from *Bifidobacterium bifidum* [39] and the recently characterized  $\alpha$ -1,3/4-L-fucosidase CpAfc2 from *Clostridium perfringens* [40] were included in the study. The former has previously been engineered into an  $\alpha$ -1,3/4-L-fucosynthase [19] and its orthologue BiAfcB from *Bifidobacterium longum* subsp. *infantis* was recently studied and engineered to attain improved transfucosylation activity [15].

Starting with the amino acid sequences listed in GenBank (Table 1), signal peptides were predicted by the SignalP 4.1 server [29] and subsequently removed from the sequences of FgFCO1, Tffuc1, BbAfcB, CpAfc2, and NixE prior to vector construction. No signal peptides were predicted in the deposited sequences of Mfuc5 or Tm $\alpha$ Fuc. In addition, presence of transmembrane regions was predicted by the TMHMM server v. 2.0 [30], and the predicted C-terminal region in BbAfcB was removed prior to vector construction (the resulting enzyme was truncated by another six amino acids compared to the truncated version expressed previously [39]). After inspection of homology models and previous expression reports [24, 35, 36, 39, 40], it was decided to place the desired His<sub>6</sub>-tag at the C-terminus of all  $\alpha$ -L-fucosidase constructs made for the current work, i.e. FgFCO1, Tffuc1, BbAfcB, CpAfc2, and NixE. In Mfuc5 and Tm $\alpha$ Fuc, which were expressed previously [14], the His<sub>6</sub>-tag was N-terminal.

BbAfcB, CpAfc2, and NixE were successfully expressed in *E. coli* BL21 (DE3). Expression levels were 38, 19, and 39 mg/L, respectively, when determined after IMAC purification and desalting by gel filtration. *E. coli* BL21 (DE3) did not grow in 400 mL-scale when transformed with vectors harbouring TfFuc1 and FgFCO1. Instead, TfFuc1 was expressed in *E. coli* BL21 (DE3) pLysS, but was only partially soluble (Figure S1b). Decreasing the temperature from 25°C to 18°C after IPTG induction increased the expression yield from 4.5 mg/L to 9 mg/L (determined after IMAC purification and desalting). The purity was assessed by SDS-PAGE, which also confirmed the expected molecular sizes (Figure S1; Table 2).

FgFCO1 was only present in the insoluble fraction when expressed in *E. coli* BL21 (DE3) pLysS (Figure S1c). Instead, FgFCO1 – being of eukaryotic origin – was expressed in *Pichia pastoris* X-33 at 35 mg/L (determined after IMAC purification and desalting). The purity was assessed by SDS-PAGE, also confirming the expected molecular size of the glycosylated enzyme as well as of the EndoH treated protein (Figure S1d; Table 2 [25]).

### Hydrolytic activity

The hydrolytic activity of the seven  $\alpha$ -L-fucosidases was quantified on the synthetic substrate *p*NP- $\alpha$ -L-Fuc, on the HMO molecules 2'FL and 3FL, as well as on the XyG fragment XXFG – a defined nonasaccharide extracted from apple (see [20] or [27] for XyG unit nomenclature). In turn, the  $\alpha$ -L-fucosidases active on XXFG were also tested on the XyG-rich fraction extracted from citrus peel (Table 3).

As expected, the GH29B  $\alpha$ -L-fucosidases were only active on the  $\alpha$ -1,3-fucosylated substrate 3FL (Table 3). Between them, BbAfcB (57 U/ $\mu$ mol) was significantly more active than CpAfc2 (17 U/ $\mu$ mol). This activity is in accordance with the linkage specificity previously reported [39, 40] and in agreement with their classification in subfamily GH29B, since these enzymes are generally only active on  $\alpha$ -1,3- and  $\alpha$ -1,4-fucosylations and show no activity on *p*NP- $\alpha$ -L-Fuc [18]. Among the GH29A  $\alpha$ -L-fucosidases, only Mfuc5 and TfFuc1 showed detectable activity on 3FL (3 U/ $\mu$ mol and 0.5 U/ $\mu$ mol, respectively).

GH29A  $\alpha$ -L-fucosidases Mfuc5, TfFuc1, Tm $\alpha$ Fuc, and NixE were all active on *p*NP- $\alpha$ -L-Fuc, which was in accordance with literature data [14, 35-37] (Table 3). In contrast, FgFCO1 showed no detectable activity on *p*NP- $\alpha$ -L-Fuc; a very low activity of FgFCO1 on *p*NP- $\alpha$ -L-Fuc has indeed been described previously [24]. Mfuc5 had more than 5 times higher specific activity (582 U/ $\mu$ mol) on *p*NP- $\alpha$ -L-Fuc than TfFuc1 (106 U/ $\mu$ mol), which in turn was twice as active as Tm $\alpha$ Fuc (47 U/ $\mu$ mol) and 5 times as active as NixE (21 U/ $\mu$ mol). Indeed, Mfuc5 had a significantly higher specific activity than the other  $\alpha$ -L-fucosidases on all the substrates (Table 3).

Focusing on the  $\alpha$ -1,2-fucosylated substrates, Mfuc5 had similar specific activity on 2'FL (37 U/ $\mu$ mol) and XXFG (34 U/ $\mu$ mol). For TfFuc1, NixE, FgFCO1 and especially for Tm $\alpha$ Fuc, the specific activity was lower on the XyG nonasaccharide XXFG as compared to the HMO 2'FL (Table 3). Among these four,

TfFuc1 was the more active on XXFG (1.7 U/ $\mu$ mol), whereas Tm $\alpha$ Fuc was the more active on 2'FL (7 U/ $\mu$ mol). Apart from NixE, which has only been reported active on *p*NP- $\alpha$ -L-Fuc as well as on an  $\alpha$ -1,3-fucosylated *N*-glycan [35], all the GH29A have previously been reported active on 2'FL [14, 25, 36]. In contrast, only FgFCO1 has been reported active on XXFG [24] and Mfuc5 on citrus XyG [14]. The XXFG preparation contained an octasaccharide impurity (possibly XXF) which was hydrolysed at the same rate as the main XXFG component, indicating that such a small difference in substrate structure and size did not influence the catalytic efficiency (Figure S2).

Since all five GH29A  $\alpha$ -L-fucosidases showed detectable activity on XXFG, their hydrolytic activity was also tested on XyG extracted from citrus peel. The XyG-rich substrate resulting from the alkaline extraction had a Fuc content of 14 mg per g dry matter corresponding to 2 mol% (Table S2). Again, Mfuc5 was significantly more active than the other  $\alpha$ -L-fucosidases with a specific activity of 8.4 U/ $\mu$ mol (Table 3), but the specific activity of Mfuc5 was 4 times lower on XyG than on XXFG. The activity decrease was even more severe for TfFuc1, NixE, and FgFCO1 – and for Tm $\alpha$ Fuc it was not possible to detect any activity on XyG at the tested enzyme concentration (0.5  $\mu$ M; Table 3). A decrease in activity on XyG compared to XXFG may be expected since the small XXFG nonasaccharide is a much more accessible substrate than the large XyG polymer where the side chains wrap tightly around the backbone [21]. The same tendency, i.e. lower activity in higher molecular weight plant cell wall xyloglucan substrates, has also been observed for uncharacterized  $\alpha$ -L-fucosidases present in plant extracts [41, 42]. Nevertheless, four of the tested GH29A  $\alpha$ -L-fucosidases were in fact hydrolytically active on citrus XyG – and all five on the XXFG nonasaccharide – thus adding to the very few GH29  $\alpha$ -L-fucosidases known to be active on XyG.

When studying the 3D structures of the active sites of the seven  $\alpha$ -L-fucosidases, it was evident that the division into GH29 subfamilies A and B is indeed linked to the active site structure [18]: Among the  $\alpha$ -L-fucosidases studied in the current work, the two GH29B members CpAfc2 and BbAfcB showed an active site structure which was similar to each other, but markedly different from that of the GH29A  $\alpha$ -L-fucosidases (Figure S3). Most of the GH29A  $\alpha$ -L-fucosidases showed similar active site structures (Figure S3 c-f); however, the opening of the Tm $\alpha$ Fuc active site was partially shielded by a loop structure (Figure S3g). This may limit its access on a large substrate such as extracted XyG.

### Thermal stability

The thermal stability of the seven  $\alpha$ -L-fucosidases was tested by incubation at 40-95°C for up to 1 h followed by assessment of the remaining hydrolytic activity on the most preferred substrate, i.e. *p*NP- $\alpha$ -L-Fuc for Mfuc5, TfFuc1, NixE, and Tm $\alpha$ Fuc, 2'FL for FgFCO1, and 3FL for BbAfcB and CpAfc2 (Table 3).

All enzymes had half-lives above 9 h at 40°C (Table 4), justifying the use of this reaction temperature throughout the work. As expected, Tm $\alpha$ Fuc being of thermophilic origin was much more

thermally stable than the other  $\alpha$ -L-fucosidases. Among those, Mfuc5, TfFuc1, and NixE were the least thermostable with half-lives of 2-4 min at 55°C. Especially Mfuc5 appeared sensitive to temperature increases with a significantly lower half-life of just 48 min at 45°C (Table 4). FgFCO1, BbAfcB, and CpAfc2 appeared more thermostable with half-lives of 2-8 h at 50°C and 28-42 min at 55°C (Table 4). CpAfc2 appeared more stable than the others at 60°C with a half-life of 31 min, which was however still 3 times lower than that of Tm $\alpha$ Fuc (97 min) (Table 4; Figure S4).

### Transfucosylation activity

#### *Transfucosylation of LNT and LNnT with GH29B $\alpha$ -L-fucosidases*

HMO core structures LNT and LNnT were used as acceptors in transfucosylation catalysed by GH29B  $\alpha$ -L-fucosidases BbAfcB and CpAfc2 with 3FL as fucosyl donor in order to investigate the potential of these two enzymes in the production of more complex HMO molecules. Both enzymes could catalyse transfucosylation at an acceptor-to-donor ratio (A:D) of 10; however, CpAfc2 was much more efficient than BbAfcB (Figure 1a). For LNT, CpAfc2 reached a molar yield of LNFP II of 39% after 10 minutes of reaction, whereas the transient maximum yield for BbAfcB was only 13% after 5 minutes (Figure 1a). Considering the fact that BbAfcB was more than 3 times as hydrolytically active on 3FL than CpAfc2 (Table 3), the transfucosylation efficiency of BbAfcB is most likely hampered by its considerable hydrolytic activity as also evidenced by the rapid decline in LNFP II concentration from 5 to 30 minutes of reaction (Figure 1a). In fact, the hydrolytic activity of BbAfcB was reported to be even higher on LNFP II and LNFP III compared to 3FL [39]; thus, severe secondary hydrolysis was expected for BbAfcB. For CpAfc2, a transient maximum was also observed; however, the yield of LNFP II was still 30% after 30 minutes, confirming a lower rate of secondary hydrolysis in this enzyme (Figure 1a). Both enzymes exhibited much lower activity when using LNnT as acceptor, but CpAfc2 was still superior in transfucosylation compared to BbAfcB: with CpAfc2 a maximum LNFP III yield of 5% was reached after 10 min, whereas it was 1.4% for BbAfcB after 5 min, both being transient maxima (Figure 1a).

For both enzymes, the expected strict  $\alpha$ -1,3/4-regioselectivity of GH29B  $\alpha$ -L-fucosidases was observed: with LNT an  $\alpha$ 1,4-linkage was formed to GlcNAc yielding LNFP II (Figure 2a) and with LNnT an  $\alpha$ -1,3-linkage was formed to GlcNAc yielding LNFP III (Figure 2b). Indeed, only  $\alpha$ -1,3/4-fucosylation of the GlcNAc residues was detected by LC-ESI-MS (Figure S6; Figure S7), underlining the potential of these enzymes for highly specific HMO production (Figure 2). Interestingly, both enzymes clearly preferred the LNT acceptor, where the Gal moiety at the non-reducing end is  $\beta$ -1,3-linked to GlcNAc and the O-4 position is available for transfucosylation, over LNnT where the opposite is the case (Figure 1a; Figure 2).

Since CpAfc2 was superior in transfucosylation, it was also tested at an A:D of 1. Previously, a yield of 6% LNFP II was reported when using the  $\alpha$ -1,3/4-L-fucosidase BiAfcB from *B. longum* subsp. *infantis* and

20 mM 3FL as donor and 20 mM LNT as acceptor [15]. When tested at the same substrate concentrations, CpAfc2 showed a 35% yield of LNFP II after 35 min when using LNT and a 1.2% yield of LNFP III after 10 min when using LNnT (Figure 1b). Thus, for LNT the yield was not significantly lowered at an A:D of 1, although the transient maximum yield was reached 20 minutes later (Figure 1b). For LNnT where CpAfc2 had lower activity, the 10 times increase in A:D did however have a positive effect, i.e. the maximum yield of LNFP III was increased 4 times (Figure 1b). When studying the L321P/F34I mutant of BiAfcB engineered for improved transufucosylation activity at the same conditions, a 17% transufucosylation yield was obtained with LNT and 21% with LNnT [15]. Hence, the acceptor preference of mutated BiAfcB appears to be the opposite of BbAfcB and CpAfcB, which both had higher activity on LNT compared to LNnT (Figure 1a).

The transufucosylation yield reported here for CpAfc2 is on the level of those reported for the best mutant of BiAfcB (L321P/F34I) at substrate concentrations of 200 mM [15]. Recently, yields as high as 60% were reached by introduction of a number of different amino acid substitutions in BiAfcB [43]. However, reaching a transufucosylation yield of 39% with WT CpAfc2 emphasizes its promising potential for future transufucosylation studies.

#### **Transufucosylation of LNT with GH29A $\alpha$ -L-fucosidases**

The  $\alpha$ -L-fucosidases belonging to subfamily GH29A are known to be less regioselective than those in GH29B [18]; thus, several reports on poor regioselectivity in transufucosylation catalysed by GH29A  $\alpha$ -L-fucosidases exist [14, 38, 44]. Indeed, for the GH29A  $\alpha$ -L-fucosidases two or three peaks corresponding to LNFP isomers were observed in the extracted ion chromatograms, indicating poor regioselectivity of these enzymes (Figure S8; Figure S9). In the current work, LNFP isomers were detected, identified, and quantified by LC-ESI-MS aided by enzymatic hydrolysis of the obtained products by  $\alpha$ -L-fucosidases Tm $\alpha$ Fuc and BbAfcB, which are specific towards  $\alpha$ -1,2-fucosyl linkages and branched  $\alpha$ -1,3/4-linkages, respectively (Table 3) [14, 19, 39]. Known HMO LNFP isomers have characteristic fragment ions which can be used for identification by MS<sup>2</sup> [45]. In the case of LNFP I which has an  $\alpha$ -1,2-fucosylation on the non-reducing end Gal moiety,  $\alpha$ -1,3-,  $\alpha$ -1,4-, or  $\alpha$ -1,6-fucosylation on this moiety would give rise to the same  $m/z$  325 fragment ion that is indicative of LNFP I. In contrast, the  $m/z$  348 fragment ion indicative of LNFP II can only arise from LNFP II itself or from an isomer where Fuc is  $\alpha$ -1,6-linked to GlcNAc. Such an isomer would not be hydrolysed by BbAfcB as LNFP II would [39], nor would it have a similar expected retention time on the PGC column since an  $\alpha$ -1,6-linkage would result in a more planar structure than the  $\alpha$ -1,4-linkage of LNFP II thus giving rise to a longer retention time of the isomer than of LNFP II [46].

The GH29A  $\alpha$ -L-fucosidases Mfuc5, NixE, FgFCO1, TfFuc1, and Tm $\alpha$ Fuc, were used for transufucosylation of LNT with 2'FL as fucosyl donor. At an A:D of 10, all five enzymes catalysed formation of LNFP I isomers reaching total molar transufucosylation yields ranging from 10% to 35% within 60 minutes of

reaction (Figure 3). In general, the specific hydrolytic activities of these enzymes towards FLs were lower than those of CpAfc2 and BbAfcB (Table 3), which may explain why no transient maximum was reached within the 60-minutes reaction time (Figure 3).

Of the five GH29A  $\alpha$ -L-fucosidases, TfFuc1 exhibited superior transfucosylation activity reaching a total molar LNFP I isomer yield of 35% after 60 min of reaction. The total LNFP I isomer yields obtained with the other four enzymes were lower, but fairly similar: 18% with NixE, 15% with FgFCO1, 12% with Mfuc5, and 10% with Tm $\alpha$ Fuc (Figure 3). The low – but yet existent – transfucosylation activity of Tm $\alpha$ Fuc has previously been described, and has been dramatically improved by protein engineering [38]. The WT Tm $\alpha$ Fuc showed poor regioselectivity when transfucosylating *p*NP- $\beta$ -Gal with *p*NP- $\alpha$ -L-Fuc as donor, whereas strict  $\alpha$ -1,2-regioselectivity was observed in the best mutant (P25) [38]. However, when this mutant Tm $\alpha$ Fuc P25 was used for transfucosylation of LNT with 2'FL as donor, the main product was Fuc- $\alpha$ 1,3-LNT (13% yield) with a minor byproduct (1% yield) [15]. For all the GH29A  $\alpha$ -L-fucosidases employed here, two or three LNFP I isomers were observed. However, the same isomer (herein called LNFP I isomer 1) was dominant for all of them, and for FgFCO1, Tm $\alpha$ Fuc, and NixE the minor peaks never exceeded a transfucosylation yield of 0.2%, i.e. their transfucosylation activity on LNT was very regioselective towards formation of LNFP I isomer 1 (Figure 3a). Only this dominant peak was degraded in the Tm $\alpha$ Fuc-catalyzed hydrolysis, indicating that it has an  $\alpha$ -1,2-linkage and is thus suggested to be LNFP I (Figure S10a-e). For TfFuc1, LNFP I isomer 1 corresponded to 93% of the total LNFP produced, but for Mfuc5 the regioselectivity was poorer: the yield of LNFP I isomer 1 was 7.3% while that of LNFP I isomer 2 was 4.2%, corresponding to 53% and 31% of the total LNFP, respectively (Figure 3b). LNFP I isomer 2 and LNFP I isomer 3 (observed only for FgFCO1 and only in trace amounts) were only partially degraded by Tm $\alpha$ Fuc after 24 h, indicating that the terminal Fuc is attached by an  $\alpha$ -1,3-,  $\alpha$ -1,4-, or  $\alpha$ -1,6-linkage (Figure S10a-e). Neither of the LNFP I isomers were degraded by BbAfcB, which is active on  $\alpha$ -1,3/4-fucosylations with a branched Gal residue and thus not expected to be active on the linear LNFP I isomers [19, 39] (Figure S10a-e).

The poor regioselectivity of the GH29A  $\alpha$ -L-fucosidases TfFuc1 and Mfuc5 was not confined to mixed linkage formation on the non-reducing Gal moiety. Indeed, an LNFP II isomer was also detected in the transfucosylation catalysed by Mfuc5 and TfFuc1 (Figure 3b). Given its long retention time on the PGC column, the fact that it was not hydrolysed by BbAfcB (Figure S10d-e), and the reported hydrolytic activity of TfFuc1 on  $\alpha$ -1,6-fucosyl linkages [36], this is likely to be an LNFP II isomer where Fuc is  $\alpha$ -1,6-linked to GlcNAc. In comparison, LNFP II produced by BbAfcB and CpAfc2 (Figure 1; Figure 2) was completely hydrolysed by BbAfcB and had a shorter retention time on the PGC column (Figure S10f-g). For Mfuc5 – the least regiospecific of the two – a 2.2% yield of this LNFP II isomer was detected, corresponding to 16% of the total LNFP produced. For TfFuc1, the yield of the LNFP II isomer obtained within 60 min was only 0.6%,

corresponding to less than 2% of the total LNFP (Figure 3b). The low regioselectivity of Mfuc5 was also observed previously [14].

#### ***Transfucosylation of lactose using XyG as donor substrate***

Searching for a natural and abundant fucosyl donor, the main aim of the current work was to assess the activity of the  $\alpha$ -L-fucosidases on XyG – and not least to determine their ability to use XyG as fucosyl donor in transfucosylation. Aiming at industrial application, lactose was used as acceptor substrate since it is abundant in dairy products and side streams. The four GH29A  $\alpha$ -L-fucosidases which had exhibited hydrolytic activity on XyG, i.e. FgFCO1, NixE, TfFuc1, and Mfuc5 were tested in transfucosylation of 100 mM lactose using the XyG-rich fraction extracted from citrus peel as donor at a concentration corresponding to 2 mM bound Fuc. The reaction was monitored for up to 24 h (Figure 4).

FgFCO1 catalysed formation of 2'FL reaching a molar yield based on the donor of 8% after 4 h and 14% after 24 h (Figure 4a). Interestingly, the rate of formation of 2'FL initially exceeded that of Fuc; however, after 24 h the concentration of released Fuc was twice as high as the 2'FL concentration (Figure 4a). No other FL isomers were detected (Figure S11). The specific hydrolytic activity of FgFCO1 on 2'FL is low (Table 3), and this appears to be reflected in the transfucosylation of lactose with XyG, where no transient maximum was observed within the 24 h of reaction (Figure 4).

As previously reported [14], the main transfucosylation product of Mfuc5 was an unidentified FL, which was neither 2'FL nor 3FL (Figure S11). The formation of the unidentified FL reached a transient maximum yield of 18% after 2 h (Figure 4d). Small amounts of 2'FL were also formed, the maximum yield being 0.6% after 3h. Compared to FgFCO1, the formation of free Fuc was much higher for Mfuc5, confirming the markedly higher hydrolytic activity of this enzyme (Table 3; Figure 4d).

For FgFCO1 and Mfuc5, the transfucosylation yields obtained with XyG and lactose (14% and 18%, respectively) were comparable to those obtained with 2'FL and LNT (15% and 14%, respectively; Figure 3). For NixE and TfFuc1, which had specific hydrolytic activities on XyG compared to that of FgFCO1 (Table 3), this was not the case. Indeed, NixE showed only hydrolytic activity on XyG (Figure 4b). Similarly, TfFuc1 showed mainly hydrolytic activity on XyG with only negligible amounts of 2'FL formed, i.e. the maximum level of 2'FL detected by HPAEC-PAD after 4 h corresponded to a molar yield on the fucosyl donor of 0.7% (Figure 4c). The product could not be detected by LC-MS (Figure S11). This result is especially striking for TfFuc1, which was superior in terms of yields in the LNT transfucosylation (Figure 3). Possibly, lactose is a poor acceptor substrate for TfFuc1. Indeed, in a preliminary transfucosylation study using 10 mM *p*NP- $\alpha$ -L-Fuc as donor and 100 mM lactose as acceptor at pH 7.0 and 45°C, TfFuc1 mainly catalysed the formation of the self-condensation product Fuc-Fuc-*p*NP (linkage unknown) and only trace amounts of 2'FL, whereas the main product of Mfuc5 was the unidentified FL with Fuc-Fuc-*p*NP as a side product. Similarly, an  $\alpha$ -L-

fucosidase from *Alcaligenes* sp. had higher transfucosylation activity with *N*-acetyllactosamine (LacNAc) as acceptor than with lactose and the same was the case for their  $\beta$ -*O*-methylated derivatives; although the fact that A:D was 25 for LacNAc and 19 for lactose may have influenced the result, the presence of the *N*-acetyl group on the reducing end GlcNAc residue still significantly improved the transfucosylation on the non-reducing end Gal residue [47]. Thus, the acceptor specificity of TfFuc1 may be more limited than usually seen for other retaining glycosidases [12]. In the homology modelled 3D structure of TfFuc1, a tunnel is observed at the acceptor binding site (Figure 5a-b); this was not observed for the other  $\alpha$ -L-fucosidases investigated here (Figure S3). When docking the lactose acceptor into the TfFuc1 3D structure, it is situated too deep inside the tunnel to function as acceptor substrate (Figure 5c). This may explain why LNT (which has lactose at the reducing end of the tetraose structure) was observed to be a markedly better acceptor for TfFuc1-catalysed transfucosylation than lactose. In contrast, docking lactose into the covalent intermediate model of FgFCO1 lead to a positioning of the 2'-OH group between the covalently bound C1 of Fuc and the general base E288 (Figure 5d).

Tm $\alpha$ Fuc was also tested for XyG transfucosylation since it had shown detectable hydrolytic activity on the XXFG nonasaccharide, but as expected from its hydrolytic activity on XyG (Table 3) and from previous results [14], neither hydrolytic nor transfucosylation activity was observed on citrus peel XyG. Using immune-glycan microarrays, some hydrolytic activity of Tm $\alpha$ Fuc was previously detected on XyG from *Arabidopsis thaliana* and *Sambucus nigra*; however, this activity was very low compared to that of Mfuc5 [14]. As hypothesized above, XyG accessibility may be limited for Tm $\alpha$ Fuc by the loop structure at the entrance to its active site (Figure S3g). As previously shown for sialidases, transglycosylation activity is not only linked to how open the active site binding pocket is, but also to specific interactions inside or near the active site [48]. Indeed, in the current work both donor and acceptor specificities of each  $\alpha$ -L-fucosidase seem to influence their transfucosylation potential significantly.

## Conclusions

In conclusion, FgFCO1 was the most promising  $\alpha$ -L-fucosidase for transfucosylation from citrus peel XyG, showing better regioselectivity leading to the formation of a true HMO structure and lower hydrolytic activity than Mfuc5. The GH29B type  $\alpha$ -L-fucosidase CpAfc2 showed promising potential in the transfucosylation of LNT using 3FL as donor substrate reaching a 39% yield of LNFP II. Using 3FL or 2'FL as donor substrates, all seven  $\alpha$ -L-fucosidases were able to catalyse transfucosylation of LNT at an A:D of 10 with varying regioselectivity. For CpAfc2, decreasing A:D to 1 did not decrease the yield significantly.

Recently,  $\alpha$ -L-fucosidase activity was discovered in a strain of *Fusarium proliferatum* able to grow on pure fucoidan. While no fucoidanase activity could be detected, significant  $\alpha$ -L-fucosidase and arylsulfatase activities were observed [49]. A GH29  $\alpha$ -L-fucosidase, which has only 25% amino acid sequence

identity to FgFCO1, was purified from this strain and was shown to catalyse transfucosylation of various *p*NP-glycosides and aliphatic alcohols using *p*NP- $\alpha$ -L-Fuc as donor substrate [50]. Being plant pathogens, the *Fusarium* sp. may prove to be a good source for discovery of  $\alpha$ -L-fucosidases active on biomass substrates, which can be used for enzymatic transfucosylation.

The perspective of the findings relates to the abundance of the produced fucosylated compounds. Depending on the blood group of the mother, 30-65% of the HMOs are fucosylated [6]. Of these, HMOs with  $\alpha$ -1,2-fucosylation are the most prevalent ones with 2'FL being the most abundant HMO [5, 6]. However,  $\alpha$ -1,2-fucosylated HMOs are absent from the milk of so-called non-secretors, i.e. mothers who lack the *SE* gene encoding the human  $\alpha$ -1,2-fucosyltransferase FUT2, although some exceptions have been encountered [51]. It is commonly accepted that 20% of women are non-secretors, even though this percentage varies slightly depending on ethnic group [5, 52, 53]. For non-secretors, the most prevalent fucosylated HMO appears to be LNFP II [6], but 3FL has also been reported as being abundant in the milk of non-secretors [5]. Thus, except for the mix of LNFP isomers produced with the GH29A  $\alpha$ -L-fucosidases and the unidentified non-HMO FL produced by Mfuc5, all the structures reported in the current transfucosylation studies on GH29A and GH29B  $\alpha$ -L-fucosidases are of interest for regioselective production as well as for functional studies.

### **Acknowledgements**

This work was supported by the Danish Council for Independent Research [grant number 5054-00046] (grant to BZ).

## Figure legends

**Figure 1.** a) Transfucosylation of LNT and LNnT with GH29B  $\alpha$ -L-fucosidases CpAfc2 (black) and BbAfcB (grey) using 3FL as fucosyl donor. Transfucosylation yields are given as % molar yields on donor substrate. Reactions took place at 40°C using 10 mM 3FL, 100 mM LNT or LNnT (A:D = 10), and an enzyme concentration of 0.5  $\mu$ M; pH was 5.5 for BbAfcB and 7.0 for CpAfc2. b) Giving highest transfucosylation yields as A:D = 10 (shown in a) and with filled symbols), CpAfc2 was also tested at A:D = 1 using 20 mM 3FL and 20 mM LNT or LNnT (open symbols). The transfucosylation products were identified and quantified by LC-ESI-MS using an external LNFP V standard. By MS<sup>2</sup> fragmentation, the single reaction product was identified as LNFP II when using LNT acceptor and as LNFP III when using LNnT acceptor (Figure S6; Figure S7; Figure 2). Standard deviations of duplicate reactions are indicated.

**Figure 2.** Transfucosylation of LNT (a) and LNnT (b) using 3FL as donor and GH29B  $\alpha$ -L-fucosidases BbAfcB and CpAfc2. Since the main activity of the  $\alpha$ -L-fucosidases is hydrolytic, free Fuc and lactose is also formed.

**Figure 3.** Transfucosylation of LNT with GH29A  $\alpha$ -L-fucosidases FgFCO1, Tm $\alpha$ Fuc (pH 4.6), NixE (pH 5.0), Mfuc5, and TfFuc1 (pH 7.0) using 2'FL as fucosyl donor. Transfucosylation yields are given as molar yields on donor substrate. Reactions took place at 40°C using 10 mM 2'FL, 100 mM LNT (A:D = 10), and an enzyme concentration of 0.5  $\mu$ M. The transfucosylation products were identified and quantified by LC-ESI-MS using an external LNFP V standard (Figure S5; Figure S8; Figure S9). For all, LNFP I isomer 1, which is suggested to be LNFP I based on enzymatic product hydrolysis data (Figure S10), was the main transfucosylation product (a-b, filled black symbols). For Mfuc5 and TfFuc1 an LNFP II isomer believed to have an  $\alpha$ -1,6-fucosylation on GlcNAc was also detected (b, filled grey symbols), but in negligible amounts. Standard deviations of duplicate reactions are indicated.

**Figure 4.** Transfucosylation from XyG to lactose. Release of Fuc and formation of FLs tested at 40°C using 2.5  $\mu$ M enzyme, 2 mM Fuc bound in XyG from citrus peel, 100 mM lactose, and a) FgFCO1 at pH 4.6, b) NixE at pH 5.0, c) TfFuc1 at pH 7.0, and d) Mfuc5 at pH 7.0 (100 mM acetate buffer for pH 4.6-5.0 and 200 mM phosphate buffer for pH 7.0). Fuc (black circles), 2'FL (grey triangles), and unidentified FL (open squares) were quantified by HPAEC-PAD. Standard deviations of duplicate samples are indicated. The unidentified FL eluted around the same time as 3FL on HPAEC-PAD, but LC-MS analysis confirmed that it was not 3FL (Figure S11) as has also been observed previously [14]. The identity of 2'FL was confirmed by LC-MS analysis (Figure S11).

**Figure 5.** The TfFuc1 tunnel (a-b) and lactose docking into TfFuc1 (c) and FgFCO1 (d). A tunnel (green spheres) going deep into the structure of TfFuc1 was observed after homology modelling for the enzyme without substrate (a). For the covalent intermediate model this tunnel (green spheres) was found to be elongated and even reaching the enzyme surface (b). The lactose molecule docked into the covalent intermediate model was bound inside the tunnel of TfFuc1 (c), whereas it was positioned with the 2'-OH group between the C1 of covalently bound Fuc and the general acid/base residue E288 in FgFCO1 (d). Colour coding: active site nucleophile D214 (TfFuc1) or D226 (FgFCO1; green), covalently bound L-Fuc which adopts  $\beta$ -conformation at this step in the double displacement mechanism (cyan), Lac (purple), general/acid base residue E288 (FgFCO1; yellow).

## References

- [1] Bode L. Human milk oligosaccharides: Every baby needs a sugar mama. *Glycobiology* 2012;22:1147–62.
- [2] Bode L, Jantscher-Krenn E. Structure-function relationships of human milk oligosaccharides. *Adv Nutr* 2012;3:383S–391S.
- [3] Garrido D, Ruiz-Moyano S, Kirmiz N, Davis JC, Totten SM, Lemay DG, et al. A novel gene cluster allows preferential utilization of fucosylated milk oligosaccharides in *Bifidobacterium longum* subsp. *longum* SC596. *Sci Rep* 2016;6:35045.
- [4] Ninonuevo MR, Park Y, Yin H, Zhang J, Ward RE, Clowers BH, et al. A strategy for annotating the human milk glycome. *J Agric Food Chem* 2006;54:7471–80.
- [5] Thurl J, Munzert M, Henker J, Boehm G, Müller-Werner B, Jelinek J, et al. Variation of human milk oligosaccharides in relation to milk groups and lactational periods. *Br J Nutr* 2010;104:1261–1271.
- [6] Kunz C, Meyer C, Collado YC, Geiger L, Garcia-Mantrana Y, Bertua-Rios Z, et al. Influence of gestational age, secretor, and Lewis blood group status on the oligosaccharide content of human milk. *J Pediatr Gastroenterol Nutr* 2017;64:789–98.
- [7] Comstock SS, Li M, Wang M, Monaco MH, Kuhlenschmidt TB, Kuhlenschmidt MS, et al. Dietary human milk oligosaccharides but not prebiotic oligosaccharides increase circulating natural killer cell and mesenteric lymph node memory T cell populations in noninfected and rotavirus-infected neonatal piglets. *J Nutr* 2017;147:1041–7.
- [8] Baumgärtner F, Seitz L, Sprenger GA, Albermann C. Construction of *Escherichia coli* strains with chromosomally integrated expression cassettes for the synthesis of 2'-fucosyllactose. *Microb Cell Fact* 2013;12:40.
- [9] Huang D, Yang K, Liu J, Xu Y, Wang Y, Wang R, et al. Metabolic engineering of *Escherichia coli* for the production of 2'-fucosyllactose and 3-fucosyllactose through modular pathway enhancement. *Metab Eng* 2017;41:23–38.
- [10] Sprenger GA, Baumgärtner F, Albermann C. Production of human milk oligosaccharides by enzymatic and whole-cell microbial biotransformations. *J Biotechnol* 2017;258:79–91.
- [11] Petschacher B, Nidetzky B. Biotechnological production of fucosylated human milk oligosaccharides: Prokaryotic fucosyltransferases and their use in biocatalytic cascades or whole cell conversion systems. *J Biotechnol* 2016;235:61–83.
- [12] Zeuner B, Jers C, Mikkelsen JD, Meyer AS. Methods for improving enzymatic trans-glycosylation for synthesis of human milk oligosaccharide biomimetics. *J Agric Food Chem* 2014;62:9615–9631.
- [13] Rodríguez-Díaz J, Carbajo RJ, Pineda-Lucena A, Monedero V, Yebra MJ. Synthesis of fucosyl-*N*-acetylglucosamine disaccharides by transfucosylation using  $\alpha$ -L-fucosidases from *Lactobacillus casei*. *Appl Environ Microbiol* 2013;79:3847–50.
- [14] Lezyk M, Jers C, Kjaerulff L, Gotfredsen CH, Mikkelsen MD, Mikkelsen JD. Novel  $\alpha$ -L-fucosidases from a soil metagenome for production of fucosylated human milk oligosaccharides. *PLoS ONE* 2016;11:e0147438.
- [15] Saumonneau A, Champion E, Peltier-Pain P, Molnar-Gabor D, Hendrickx J, Tran V, et al. Design of an  $\alpha$ -L-transfucosidase for the synthesis of fucosylated HMOs. *Glycobiology* 2016;26:261–9.
- [16] Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 2014;42:D490–5.
- [17] Bissaro B, Monsan P, Fauré R, O'Donohue MJ. Glycosynthesis in a waterworld: new insight into the molecular basis of transglycosylation in retaining glycoside hydrolases. *Biochem J* 2015;467:17–35.
- [18] Sakurama H, Tsutsumi E, Ashida H, Katayama T, Yamamoto K, Kumagai H. Differences in the substrate specificities and active-site structures of two  $\alpha$ -L-fucosidases (glycoside hydrolase family 29) from *Bacteroides thetaiotaomicron*. *Biosci Biotechnol Biochem* 2012;76:1022–24.
- [19] Sakurama H, Fushinobu S, Hidaka M, Yoshida E, Honda Y, Ashida H, et al. 1,3-1,4- $\alpha$ -L-fucosynthase that specifically introduces Lewis a/x antigens into type-1/2 chains. *J Biol Chem* 2012;287:16709–19.
- [20] Schultink A, Liu L, Zhu L, Pauly M. Structural diversity and function of xyloglucan sidechain substituents. *Plants* 2014;3:526–42.

- [21] Fry SC. The structure and functions of xyloglucan. *J Exp Bot* 1989;40:1-11.
- [22] Günl M, Neumetzler L, Kraemer F, de Souza A, Schultink A, Pena M, et al. AXy8 encodes an  $\alpha$ -fucosidase, underscoring the importance of apoplastic metabolism on the fine structure of *Arabidopsis* cell wall polysaccharides. *Plant Cell* 2011;23:4025-40.
- [23] Ishimizu T, Hashimoto C, Takeda R, Fujii K, Hase S. A novel  $\alpha$ 1,2-L-fucosidase acting on xyloglucan oligosaccharides is associated with endo- $\beta$ -mannosidase. *J Biochem* 2007;142:721-9.
- [24] Paper JM, Scott-Craig JS, Cavalier D, Faik A, Wiemels RE, Borrusch MS, et al.  $\alpha$ -Fucosidases with different substrate specificities from two species of *Fusarium*. *Appl Microbiol Biotechnol* 2013;97:5371-80.
- [25] Cao H, Walton JD, Brumm P, Phillips GN. Structure and substrate specificity of a eukaryotic fucosidase from *Fusarium graminearum*. *J Biol Chem* 2014;289:25624-38.
- [26] Sulzenbacher G, Bignon C, Nishimura T, Tarling CA, Withers SG, Henrissat B, et al. Crystal structure of *Thermotoga maritima*  $\alpha$ -L-fucosidase. *J Biol Chem* 2004;279:13119-28.
- [27] Pauly M, Qin Q, Greene H, Albersheim P, Darvill A, York WS. Changes in the structure of xyloglucan during cell elongation. *Planta* 2001;212:842-50.
- [28] Ravn HC, Meyer AS. Chelating agents improve enzymatic solubilization of pectinaceous co-processing streams. *Process Biochem* 2014;49:250-7.
- [29] Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 2011;8:785-6.
- [30] Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J Mol Biol* 2001;305:567-80.
- [31] Nielsen AVF, Nyffenegger C, Meyer AS. Performance of Microbial Phytases for Gastric Inositol Phosphate Degradation. *J Agric Food Chem* 2015 63:943–950.
- [32] Krieger E, Vriend G. YASARA View—molecular graphics for all devices—from smartphones to workstations *Bioinformatics* 2014;30:2981–2982.
- [33] Krieger E, Darden T, Nabuurs SB, Finkelstein A, Vriend G. Making optimal use of empirical energy functions: Force-field parameterization in crystal space. *Proteins: Structure, Function, and Bioinformatics* 2004;57:678–683.
- [34] Chovancova E, Pavelka A, Benes P, Strnad O, Brezovsky J, Kozlikova B, et al. CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput. Biol.* 2012;8:e1002708.
- [35] Dupoirion S, Zischek C, Ligat L, Carbonne J, Boulanger A, de Bernonville TD, et al. The N-glycan cluster from *Xanthomonas campestris* pv. *campestris*. *J Biol Chem* 2015;290:6022-36.
- [36] Megson ZA, Koerdt A, Schuster H, Ludwig R, Janesch B, Frey A, et al. Characterization of an  $\alpha$ -L-fucosidases from the periodontal pathogen *Tannerella forsythia*. *Virulence* 2015;6:282-92.
- [37] Tarling CA, He S, Sulzenbacher G, Bignon C, Bourne Y, Henrissat B, et al. Identification of the catalytic nucleophile of the family 29  $\alpha$ -L-fucosidase from *Thermotoga maritima* through trapping of a covalent glycosyl-enzyme intermediate and mutagenesis. *J Biol Chem* 2003;278:47394-9.
- [38] Osanjo G, Dion M, Drone J, Solleux C, Tran V, Rabiller C, et al. Directed evolution of the  $\alpha$ -L-fucosidase from *Thermotoga maritima* into an  $\alpha$ -L-transfucosidase. *Biochem* 2007;46:1022-1033.
- [39] Ashida H, Miyake A, Kiyohara M, Wada J, Yoshida E, Kumagai H, et al. Two distinct  $\alpha$ -L-fucosidases from *Bifidobacterium bifidum* are essential for the utilization of fucosylated milk oligosaccharides and glycoconjugates. *Glycobiology* 2009;19:1010-7.
- [40] Fan S, Zhang H, Chen X, Lu L, Xu L, Xiao M. Cloning, characterization, and production of three  $\alpha$ -L-fucosidases from *Clostridium perfringens* ATCC 13124. *J Basic Microbiol* 2016;56:347-57.
- [41] O'Neill RA, White AR, York WS, Darvill AG, Albersheim P. A gas chromatographic-mass spectrometric assay for glycosylases. *Phytochemistry* 1988;27:329-33.
- [42] Farkas V, Hanna R, Maclachlan G. Xyloglucan oligosaccharide  $\alpha$ -L-fucosidase activity from growing pea stems and germinating nasturtium seeds. *Phytochemistry* 1991;30:3203-7.
- [43] Champion E, Vogel A, Bartsch S, Dekany G. Mutated fucosidase. Patent WO2016/063261 A1. April 28, 2016.

- [44] Benesová E, Lipovová P, Krejzová J, Kovalová T, Buchtová P, Spiwok V, et al. Alpha-L-fucosidase isoenzyme iso2 from *Paenibacillus thiaminolyticus*. *BMC Biotechnol* 2015;15:36.
- [45] Pfenninger A, Karas M, Finke B, Stahl B. Structural analysis of underivatized neutral human milk oligosaccharides in the negative ion mode by nano-electrospray MS<sup>n</sup> (Part 1: Methodology). *J Am Soc Mass Spectrom* 2002;13:1331–40.
- [46] Westphal Y, Schols HA, Voragen AGJ, Gruppen H. Introducing porous graphitized carbon liquid chromatography with evaporative light scattering and mass spectrometry detection into cell wall oligosaccharide analysis. *J Chromatogr A* 2010;1217:689-95.
- [47] Murata T, Morimoto S, Zeng X, Watanabe S, Usui T. Enzymatic synthesis of  $\alpha$ -L-fucosyl-N-acetyllactosamines and 3'-O- $\alpha$ -L-fucosyllactose utilizing  $\alpha$ -L-fucosidases. *Carbohydr Res* 1999;320,192-9.
- [48] Nordvang RT, Nyffenegger C, Holck J, Jers C, Zeuner B, Sundekilde UK, et al. It all starts with a sandwich: Identification of sialidases with trans-glycosylation activity. *PLOS One* 2016;11:e0158434.
- [49] Shvetsova SV, Zhurishkina EV, Bobrov KS, Ronzhina NL, Lapina IM, Ivanen DR, et al. The novel strain *Fusarium proliferatum* LE1 (RCAM02409) produces  $\alpha$ -L-fucosidase and arylsulfatase during the growth on fucoidan. *J Basic Microbiol* 2015;55:471-9.
- [50] Shvetsova SV, Shabalin KA, Bobrov KS, Ivanen DR, Ustyuzhanina NE, Krylov VB, et al. Characterization of a new  $\alpha$ -L-fucosidase isolated from *Fusarium proliferatum* LE1 that is regioselective to  $\alpha$ -(1-4)-L-fucosidic linkage in the hydrolysis of  $\alpha$ -L-fucobiosides. *Biochimie* 2017;132:54-65.
- [51] Newburg DS, Ruiz-Palacios GM, Morrow AL. Human milk glycans protect infants against enteric pathogens. *Annu Rev Nutr* 2005;25:37-58.
- [52] Ishitoya S, Yamamoto S, Mitsumori K, Ogawa O, Terai A. Non-secretor status is associated with female acute uncomplicated pyelonephritis. *BJU Int* 2002;89:851-854.
- [53] Jaff MS. Higher frequency of secretor phenotype in O blood group – its benefits in prevention and/or treatment of some diseases. *Int J Nanomed* 2010;5:901-905.
- [54] Alva V, Nam SZ, Söding J, Lupas AN. The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. *Nucleic Acids Res* 2016;44:W410-W415.
- [55] Ferrè F, Clote P. DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification. *Nucleic Acids Res* 2006;34:W182-5.

## Tables

**Table 1.** Microbial origin, GenBank accession numbers, vectors, and expression strains used for production of  $\alpha$ -L-fucosidases in the current work.

Enzyme name	Organism	GenBank no.	Vector	Features <sup>1</sup>	AAs inserted <sup>2</sup>	Expression strain
FgFCO1	<i>Fusarium graminearum</i> PH-1	AFR68935.1	pPICZ $\alpha$ A ( <i>EcoRI</i> , <i>XbaI</i> )	His <sub>6</sub> (C), Myc (C), $\alpha$ -factor, Zeo <sup>R</sup>	26-609 (end)	<i>P. pastoris</i> X-33
FgFCO1	<i>Fusarium graminearum</i> PH-1	AFR68935.1	pET22b(+) ( <i>NcoI</i> , <i>XhoI</i> ) <sup>3</sup>	His <sub>6</sub> (C), <i>pelB</i> , Amp <sup>R</sup>	26-609 (end)	<i>E. coli</i> BL21 (DE3) pLysS
TfFuc1	<i>Tannerella forsythia</i> ATCC 43037	AEW21393.1	pET22b(+) ( <i>NcoI</i> , <i>XhoI</i> ) <sup>3</sup>	His <sub>6</sub> (C), <i>pelB</i> , Amp <sup>R</sup>	21-446 (end)	<i>E. coli</i> BL21 (DE3) pLysS <sup>4</sup>
BbAfcB	<i>Bifidobacterium bifidum</i> JCM 1254	BAH80310.1	pET22b(+) ( <i>NdeI</i> , <i>XhoI</i> )	His <sub>6</sub> (C), Amp <sup>R</sup>	38-1462	<i>E. coli</i> BL21 (DE3)
CpAfc2	<i>Clostridium perfringens</i> ATCC 13124	ABG83106.1	pET22b(+) ( <i>NcoI</i> , <i>XhoI</i> ) <sup>3</sup>	His <sub>6</sub> (C), <i>pelB</i> , Amp <sup>R</sup>	26-932 (end)	<i>E. coli</i> BL21 (DE3)
NixE	<i>Xanthomonas campestris</i> pv. <i>campestris</i> ATCC 33913	AAM42160.1	pET22b(+) ( <i>NdeI</i> , <i>XhoI</i> )	His <sub>6</sub> (C), Amp <sup>R</sup>	40-561 (end)	<i>E. coli</i> BL21 (DE3)
Mfuc5 <sup>5</sup>	-	AIC77302.1	pETM-10 ( <i>NcoI</i> , <i>KpnI</i> )	His <sub>6</sub> (N), Kan <sup>R</sup>	1-419 (end)	<i>E. coli</i> C41 (DE3)
TmaFuc <sup>5</sup>	<i>Thermotoga maritima</i> MSB8	AAD35394.1	pETM-10 ( <i>NcoI</i> , <i>KpnI</i> )	His <sub>6</sub> (N), Kan <sup>R</sup>	1-449 (end)	<i>E. coli</i> BL21 (DE3) pLysS <sup>4</sup>

<sup>1</sup>Disulfide bridges were predicted via the DiANNA 1.1 software tool [54] and inspection of 3D structures (from PDB or homology models made with HHPred and MODELLER [55]). In case of clear evidence of disulfide bridges, a *pelB* leader sequence was included (using *NcoI*), except for BbAfcB, which was previously successfully expressed without [39]. N or C designates the placement of the His<sub>6</sub>-tag. <sup>2</sup>Signal peptides and transmembrane regions were predicted using the SignalP 4.1 server [29] and the TMHMM server v. 2.0 [30]. All predicted signal peptides were removed from the sequence prior to insertion into the vector. <sup>3</sup>An extra Gly at the N-terminal resulted from adding additional nucleotides to keep the reading frame. <sup>4</sup>Expression was first tested in *E. coli* BL21 (DE3), but turned out to be significantly improved in *E. coli* BL21(DE3) pLysS. <sup>5</sup>Expression done and described previously [14].

**Table 2.** GH29 subfamily and properties (optimal conditions, size, and quaternary structure) reported in the literature for the  $\alpha$ -L-fucosidases [14, 24, 25, 35-37, 39, 40], along with the apparent size of the expressed  $\alpha$ -L-fucosidases as determined by SDS-PAGE (Figure S1) and their expression yields.

	Subfamily <sub>1</sub>	pH <sub>opt</sub>	T <sub>opt</sub>	Size	Apparent size	Expression yield <sup>2</sup>
<b>FgFCO1</b>	A	4.6	-	66 kDa <sup>3</sup>	~70 kDa (75 kDa glycosylated)	35 mg/L
<b>TfFuc1</b>	A	9 <sup>4</sup>	-	51 kDa	~48 kDa	9 mg/L
<b>BbAfcB</b>	B	5.5-6	45°C	160 kDa	~150 kDa	38 mg/L
<b>CpAfc2</b>	B	7	40°C	101 kDa	~100 kDa	19 mg/L
<b>NixE</b>	A	5	37°C	58 kDa	~55 kDa	39 mg/L
<b>Mfuc5</b>	A	7	~40°C	48 kDa	~50 kDa <sup>5</sup>	-
<b>Tm<math>\alpha</math>Fuc</b>	A	4.5	≥60°C	52 kDa	~50 kDa <sup>5</sup>	-

<sup>1</sup>The subfamily division was first proposed by Sakurama and co-workers [18]; the designation here is derived from the phylogenetic tree presented in [14] and the multiple sequence alignment in [40]. <sup>2</sup>Determined after IMAC purification and desalting. <sup>3</sup>Apparent size of deglycosylated enzyme; based on SDS-PAGE the size was approx. 70 kDa with glycosylation (produced in *P. pastoris*) [25]. <sup>4</sup>Optimal pH reported to be 9.0 [36], but in the current work TfFuc1 had similar hydrolytic activities on pNP- $\alpha$ -L-Fuc at pH 7.0 and pH 9.0, with indications of inactivation over time at pH 9.0 (data not shown). <sup>5</sup>Data from [14].

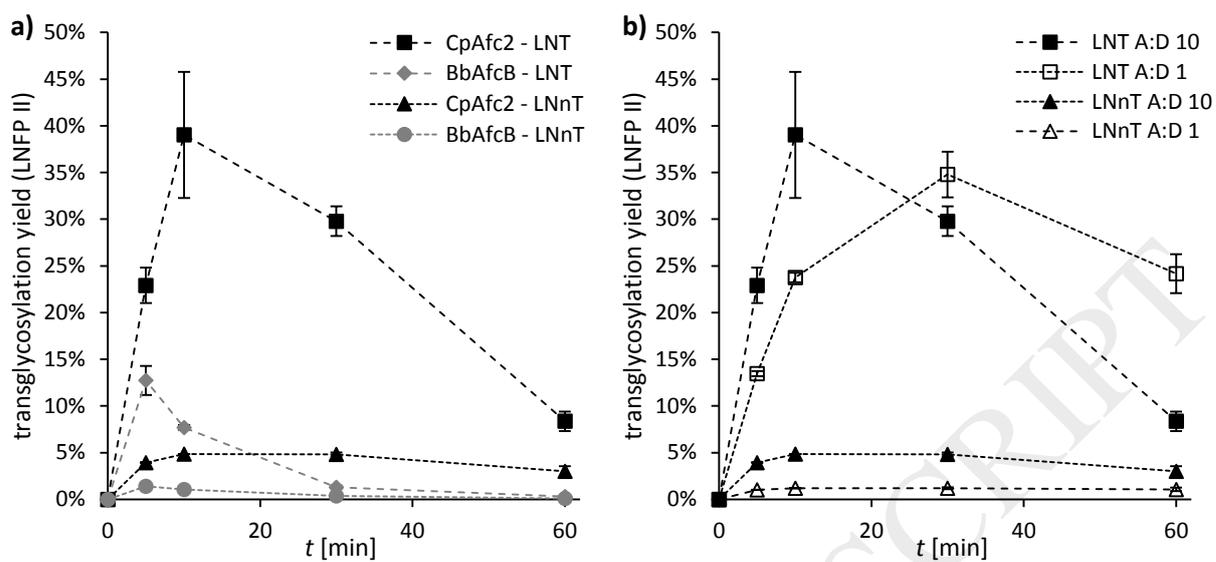
**Table 3.** Hydrolytic activity of the expressed  $\alpha$ -L-fucosidases. Activity given as  $\mu$ mol product released per minute (U) per  $\mu$ mol of enzyme for all substrates. The reaction temperature was 40°C and pH was 5.0 for NixE, FgFCO1, Tm $\alpha$ Fuc, and BbAfcB, and 7.0 for Mfuc5, TfFuc1, and CpAfc2. Only enzymes active on XXFG were tested on the XyG-rich fraction from citrus peel. Standard deviations of duplicate reactions are indicated. The GH29 subfamily is indicated since it relates to substrate specificity. Superscript letters indicate significant difference between enzymes for each substrate.

	Subfamily	pNP- $\alpha$ -L-Fuc	XXFG	XyG	2'FL	3FL
<b>Mfuc5</b>	A	582 $\pm$ 51 <sup>a</sup>	34.1 $\pm$ 4.5 <sup>a</sup>	8.4 $\pm$ 0.02 <sup>a</sup>	37.0 $\pm$ 2.8 <sup>a</sup>	2.8 $\pm$ 0.5 <sup>c</sup>
<b>TfFuc1</b>	A	106 $\pm$ 1.3 <sup>b</sup>	1.7 $\pm$ 0.1 <sup>b</sup>	0.03 $\pm$ 0 <sup>b</sup>	3.0 $\pm$ 0.5 <sup>c</sup>	0.5 $\pm$ 0.1 <sup>d</sup>
<b>NixE</b>	A	21.2 $\pm$ 0.5 <sup>c</sup>	0.5 $\pm$ 0.03 <sup>b</sup>	0.01 $\pm$ 0.002 <sup>b</sup>	3.7 $\pm$ 0 <sup>bc</sup>	0
<b>FgFCO1</b>	A	0	0.3 $\pm$ 0.02 <sup>b</sup>	0.03 $\pm$ 0.002 <sup>b</sup>	1.1 $\pm$ 0.3 <sup>c</sup>	0
<b>Tm<math>\alpha</math>Fuc</b>	A	47.3 $\pm$ 5.1 <sup>bc</sup>	0.1 $\pm$ 0.01 <sup>b</sup>	0	7.0 $\pm$ 0.5 <sup>b</sup>	0
<b>BbAfcB</b>	B	0	0	-	0	57.0 $\pm$ 1.4 <sup>a</sup>
<b>CpAfc2</b>	B	0	0	-	0	17.2 $\pm$ 0.2 <sup>b</sup>

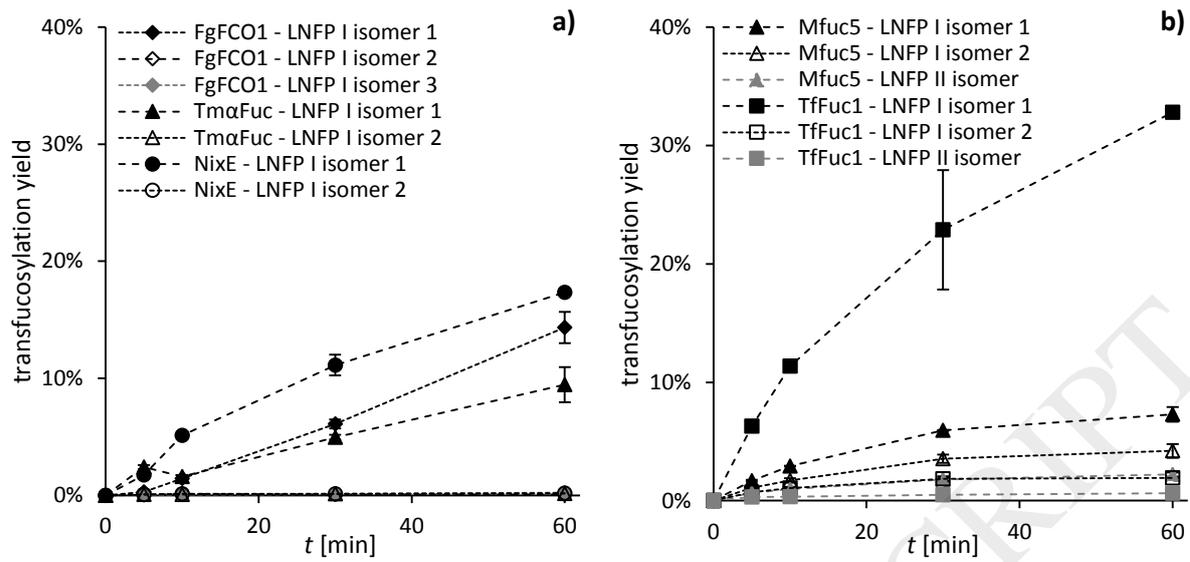
**Table 4.** Thermal stability of the expressed  $\alpha$ -L-fucosidases. Enzyme samples were incubated at 40-95°C for up to 1 h and the residual activity was measured at standard conditions (Figure S4). Half-lives (t<sub>1/2</sub>) are given

in min  $\pm$  standard deviations of duplicate reactions. Empty fields indicate that the enzymes were not tested at the given temperature, whereas the hyphen indicates that the first-order thermal inactivation constant  $k_D = 0$  (Figure S4). Superscript letters indicate significant difference between half-lives of the enzymes at each temperature ( $p < 0.05$ ).

Incubation temperature	Mfuc5	TfFuc1	NixE	FgFCO1	BbAfcB	CpAfc2	Tm $\alpha$ Fuc
40°C	790 $\pm$ 257 <sup>b</sup>	-	573 $\pm$ 86 <sup>b</sup>	621 $\pm$ 174 <sup>b</sup>	>1740 <sup>a</sup>	780 $\pm$ 17 <sup>b</sup>	-
45°C	48 $\pm$ 1.1 <sup>c</sup>	355 $\pm$ 84 <sup>ab</sup>	419 $\pm$ 67 <sup>ab</sup>	455 $\pm$ 238 <sup>a</sup>	556 $\pm$ 259 <sup>a</sup>	171 $\pm$ 14 <sup>bc</sup>	
50°C	6.2 $\pm$ 0.02 <sup>e</sup>	14 $\pm$ 0.9 <sup>e</sup>	24 $\pm$ 2.3 <sup>de</sup>	244 $\pm$ 15 <sup>b</sup>	493 $\pm$ 162 <sup>a</sup>	113 $\pm$ 14 <sup>cd</sup>	134 $\pm$ 27 <sup>c</sup>
55°C	1.7 $\pm$ 0.02 <sup>c</sup>	3 $\pm$ 0.05 <sup>c</sup>	4.3 $\pm$ 0.05 <sup>c</sup>	39 $\pm$ 2.5 <sup>a</sup>	28 $\pm$ 1.6 <sup>b</sup>	42 $\pm$ 6.3 <sup>a</sup>	
60°C	0.7 $\pm$ 0.01 <sup>c</sup>	1.6 $\pm$ 0.01 <sup>c</sup>	2.2 $\pm$ 0.02 <sup>c</sup>	6.0 $\pm$ 0.4 <sup>c</sup>	4.7 $\pm$ 0.08 <sup>c</sup>	30.6 $\pm$ 3.4 <sup>b</sup>	97 $\pm$ 28 <sup>a</sup>
70°C							28 $\pm$ 3.4
80°C							19 $\pm$ 2.9
90°C							15 $\pm$ 0.7
95°C							4.4 $\pm$ 0.03

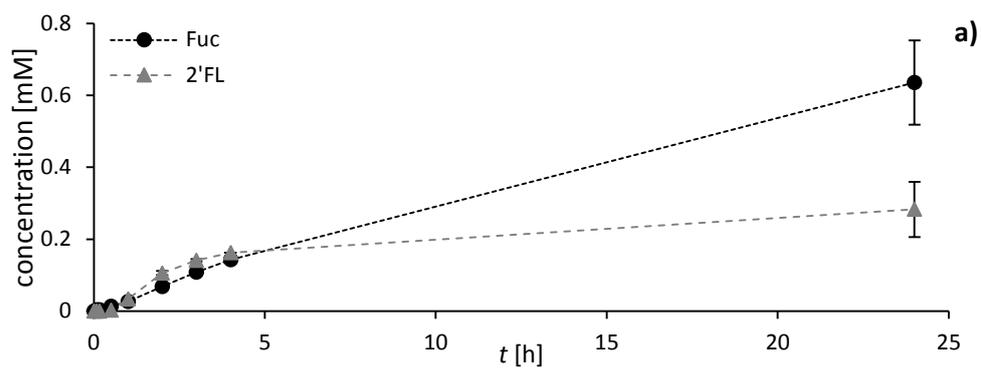


Figr-1

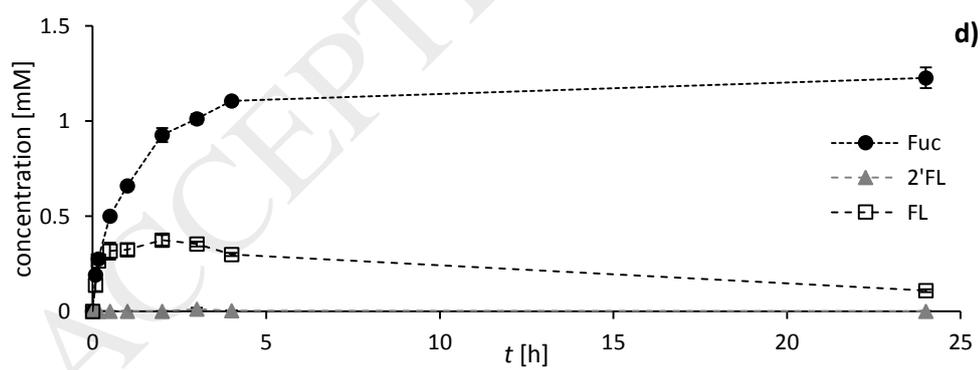
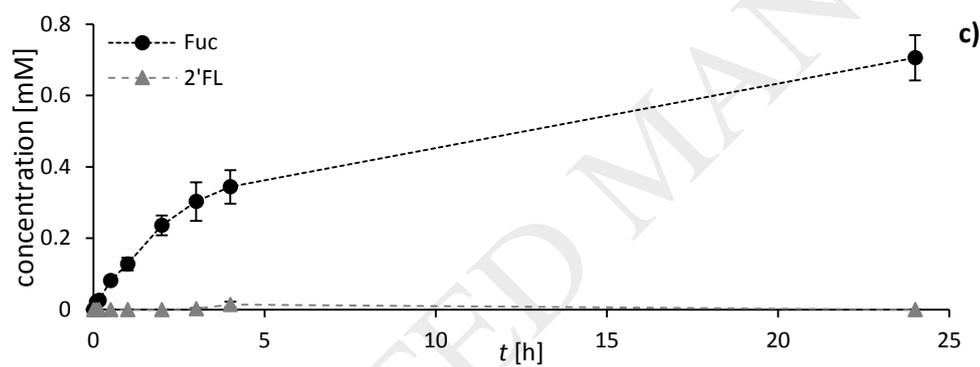
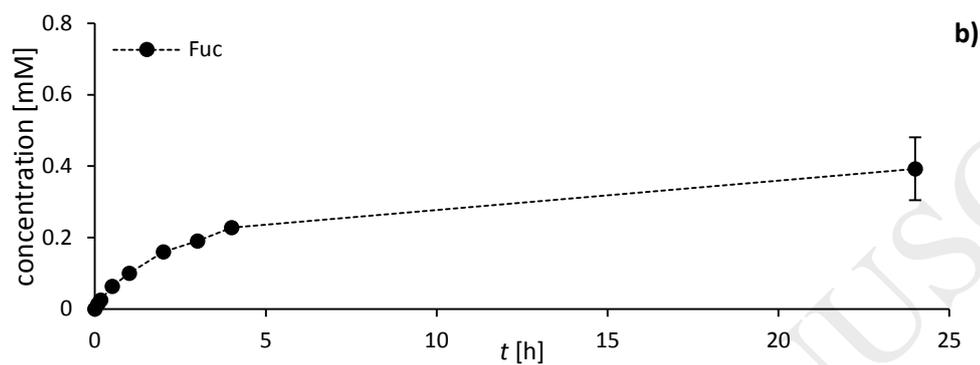


Figr-2

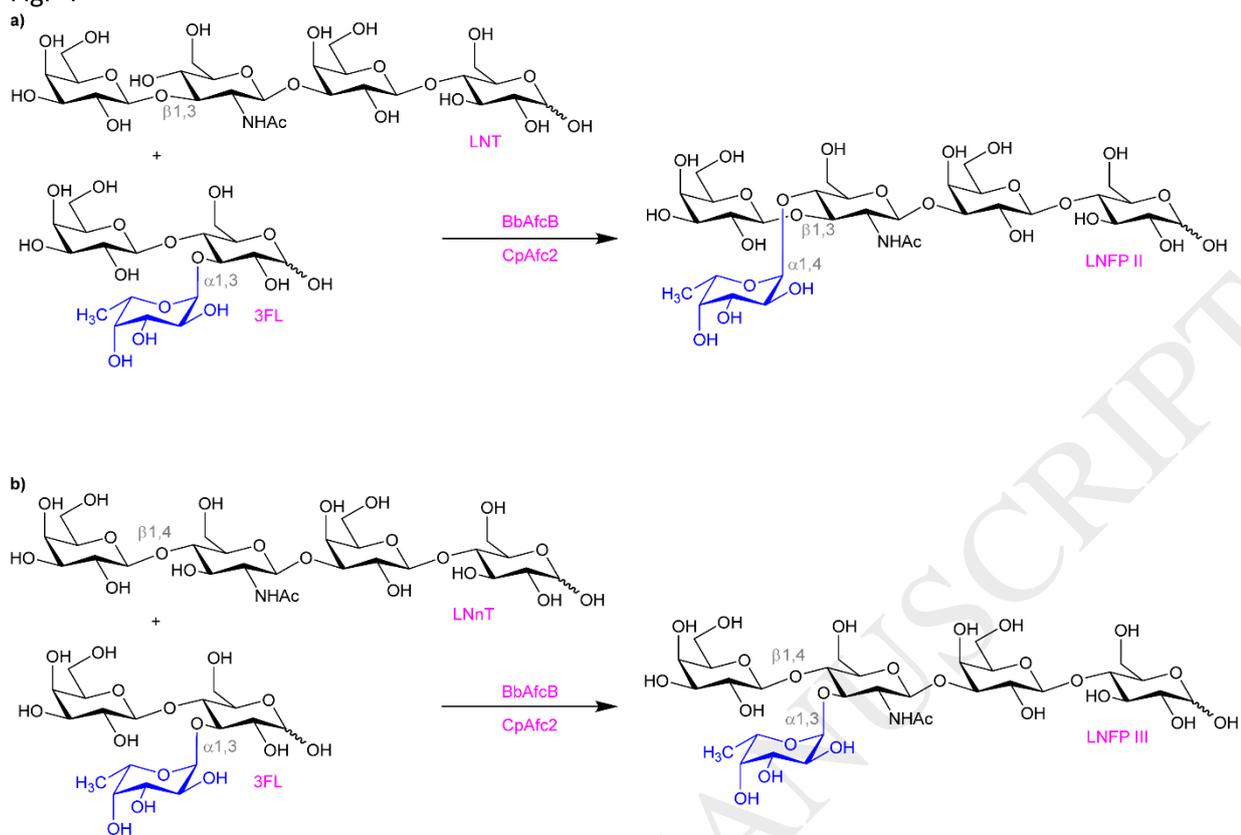
ACCEPTED MANUSCRIPT



Figr-3



Figr-4



Figr-5

