University of Dundee

Exploring the sequence-function space of microbial fucosidases

Gascuena, Ana Martínez; Wu, Haiyang; Owen, David; Hernando, Pedro J; Monaco, Serena; Penner, Matthew

Link to publication in Discovery Research Portal

Citation for published version (APA):
Gascuena, A. M., Wu, H., Owen, D., Hernando, P. J., Monaco, S., Penner, M., Gall, G. L., Gardner, R., Ndeh, D., Urbanowicz, P., Spencer, D. I. R., Walsh, M., Angulo, J., & Juge, N. (2023). *Exploring the sequence-function space of microbial fucosidases*. Research Square. https://doi.org/10.21203/rs.3.rs-3101218/v1

Exploring the sequence-function space of microbial fucosidases

Ana Martínez Gascueña

Quadram Institute Bioscience   https://orcid.org/0000-0002-6245-5823

Haiyang Wu

Quadram Institute Bioscience

David Owen

Diamond Light Source Ltd

Pedro Hernando

Quadram Institute Bioscience   https://orcid.org/0009-0003-8173-4048

Serena Monaco

University of East Anglia

Matthew Penner

Diamond Light Source Ltd

Gwenaelle Le Gall

University of East Anglia   https://orcid.org/0000-0002-1379-2196

Richard Gardner

Ludger (United Kingdom)   https://orcid.org/0000-0001-7601-7544

Didier Ndeh

University of Dundee

Paulina Urbanowicz

Ludger   https://orcid.org/0000-0003-4158-0462

Daniel Spencer

Ludger Ltd   https://orcid.org/0000-0001-6386-0890

Martin Walsh

Diamond Light Source   https://orcid.org/0000-0001-5683-1151

Jesús Angulo

Institute for Chemical Research (CSIC-University of Seville)   https://orcid.org/0000-0001-7250-5639

Nathalie Juge (✉ nathalie.juge@quadram.ac.uk)

Quadram Institute Bioscience   https://orcid.org/0000-0001-8515-1315

Abstract

Microbial α-l-fucosidases catalyse the hydrolysis of terminal α-l-fucosidic linkages with diverse substrate/linkage specificities and can be used in transglycosylation reactions to synthesise oligosaccharides. Based on sequence identity, α-l-fucosidases have been classified in distinct glycoside hydrolases (GHs) families in the carbohydrate-active enzymes (CAZy) database. Here, we explored the sequence-function space of fucosidases from GH29 family. Based on sequence similarity network (SSN) analyses, 16 GH29 α-l-fucosidases were selected for functional characterisation. Using activity assays combined with HPAEC-PAD and LC-FD-MS/MS analyses, we determined the substrate and linkage specificities of these enzymes against a range of defined oligosaccharides and glycoconjugates, revealing a range of specificities for α1,2, α1,3, α1,4 and α1,6 linked fucosylated ligands. The structural basis for the substrate specificity of GH29 fucosidase from *Bifidobacterium asteroides* towards α1-6 linkages and FA2G2 *N*-glycan was further determined by X-ray crystallography and saturation transfer difference NMR. TLC combined with electrospray ionization − MS and NMR confirmed the capacity of this enzyme to carry out transfucosylation reactions with GlcNAc and Fuc1,3GlcNAc as acceptors. Taken together, these experimental data validate the use of SSN as a reliable bioinformatics approach to predict the substrate specificity and transfucosylation activity of GH29 fucosidases.

Introduction

Carbohydrate-active enzymes (CAZymes) are responsible for the synthesis, breakdown and modification of all carbohydrates on earth. In the sequence-based classification database (ww.cazy.org), CAZymes are grouped into families covering enzymes with common folds and enzymatic mechanisms but different substrate specificities[1]. The number of CAZymes and their families is continuously expanding with glycoside hydrolases (GHs) showing an exponential increase driven largely by high-throughput microbial whole-genome and metagenomic sequencing[2].

This is for example the case of α-l-fucosidases which are classified into GH29, GH95, GH139, GH141, and GH151 families, a majority of which are from microbial sources[3]. Reflecting the high diversity of naturally occurring fucosylated structures, these enzymes show a wide range of substrate and linkage specificity cleaving the nonreducing terminal α-l-fucose (Fuc) and have numerous biological roles and applications in health and biotechnology[4].

GH95 enzymes functionally characterized so far show strict substrate specificity to the terminal Fuc α1-2Gal linkage and hydrolyse the linkage via an inverting mechanism. The GH139 and GH141 families include one functionally characterised fucosidase targeting specific α-l-fucose motifs in pectin[5]. The GH151 family includes one characterised fucosidase targeting α 1,2/3/4/6 fucosylated disaccharides[6] and two other GH151 members which remain to be functionally characterised[7,8].

In contrast, fucosidases from the GH29 family have been extensively studied and are reported to act on a wide range of substrates with hydrolysis proceeding via a retaining mechanism. This family covers fucosidases with substrate specificities against Fuc(α1,2/3/4/6) motifs. Some of the GH29 fucosidases have relaxed substrate specificities and can act on 4-nitrophenyl α-l-fucopyranoside (pNP-Fuc) (EC 3.2.1.51), while other

fucosidases show strict specificity for terminal α-(1–3/4)-fucosidic linkages with little/no activity on pNP-Fuc (EC 3.2.1.111), which led to an attempt to subdivide the GH29 family into substrate-based specificity GH29-A and GH29-B, but more accurate classification is needed[9]. We previously reported the substrate and linkage specificities of fucosidases from the human gut symbiont *Ruminococcus gnavus*, revealing a GH29 fucosidase with the capacity to recognize sialic acid-terminated fucosylated glycans (sialyl Lewis X/A epitopes) and hydrolyze α1–3/4 fucosyl linkages in these substrates without the need to remove sialic acid[10]. In addition, GH29 fucosidases are increasingly being considered as glyco-tools for their capacity to synthesise oligosaccharides by transglycosylation, as reported for AlfB and AlfC from *Lactobacillus casei* BL23[11,12]. Given the wide enzymatic diversity within the GH29 family, there is great interest in mining this family for applications[13,14] and several bioinformatics-based approaches are being tested to better predict substrate specificity and transglycosylation ability of these enzymes[15,16].

Here we used sequence similarity network (SSN) to explore the microbial GH29 sequence-function space and experimentally validated the substrate specificity and transfucosylation capacity of the novel enzymes identified by SSN.

Materials and Methods
Materials

All chemicals were obtained from Sigma (St Louise, MO, USA) unless otherwise stated. 2'-fucosyllactose (2'FL), 3-fucosyllactose (3FL) and difucosyllactose (DFL) were obtained from Glycom/DSM (Esbjerg, Denmark). Blood group A type II (BgA), Blood group B type II (BgB), Blood group H type II (BgH) and LewisY (LeY) were obtained from Elicityl (Crolles, France). Lewis A trisaccharide (LeA), 3'-sialyl Lewis A (sLeA), Lewis X trisaccharide (LeX), 3'-sialyl Lewis X (sLeX), 2-acetamido-2-deoxy-6-O-(a-l-fucopyranosyl)-D-glucopyranose (Fuc1,6GlcNAc), 2-acetamido-2-deoxy-4-O-(α -L-fucopyranosyl)-D-glucopyranose (Fuc1,4GlcNAc), 2-acetamido-2-deoxy-3-O-(α -L-fucopyranosyl)-D-glucopyranose (Fuc1,3GlcNAc), 4-nitrophenyl α -l-fucopyranoside (pNP-Fuc), 2-Chloro-4-nitrophenyl-α -l-fucopyranoside (CNP-Fuc), 2-Chloro-4-nitrophenol (CNP) and N-acetyllactosamine (LacNAc) were obtained from Biosynth Ltd (Compton, UK). FA2G2 N-glycan was from Ludger (Oxford, UK). IgG was purified from human serum using the protein A IgG purification kit from Thermofisher (Carlsbad, US). PNGase B035DRAFT_03341[17] was a kind gift from Dr Lucy Crouch (Newcastle University). Phospholipase A2 (PLA2) from honeybee venom (*Apis mellifera*) was purchased from Sigma (St Louise, MO, USA). Recombinant fucosidases E1_10125 from *R. gnavus* E1 and ATCC_03833 from *R. gnavus* ATCC 29149 were produced in-house as previously reported[10].

Bioinformatics analyses

For sequence similarity networks (SSN) analysis, the sequences encoding GH29 fucosidases were extracted from CAZy database (www.cazy.org). A total of 9505 GH29 sequences from CAZy (last update 2022-10-18) were winnowed down to 2971 sequences following a sequence identity cut-off at 0.8 via CD-HIT suite[18]. The amino acid sequences were then used to generate SSN using the Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST) with an alignment score threshold of 96 (40% sequence identity)[19]. The SSN was visualised using Cytoscape 3.9.1.

Cloning, expression and purification of fucosidases

The GH29-encoding genes were synthesised exempt of the signal peptide sequence and cloned into pET28a with N terminal His_6-tag by Prozomix (Haltwhistle, UK). TT1819D218A and TT1819D218N mutants were synthesised by NZYTech (Lisboa, Portugal). *Escherichia coli* TunerDE3 pLacI cells were transformed with the recombinant plasmids according to manufacturer's instructions. Expression was carried out in 1L LB media growing cells at 37°C until OD_{600} reached 0.3 to 0.6 and then induced at 16°C for 20–22 h. The cells were harvested by centrifugation at 4,000g for 35 min. The His-tagged proteins were purified by immobilized metal affinity chromatography (IMAC) and further purified by gel filtration on an ÄKTApure (Cytiva, Little Chalfont, UK). Protein purification was assessed by standard SDS−polyacrylamide gel electrophoresis using the NuPAGE Novex 4–12% Bis-Tris (Life Technologies, Paisley, UK). Protein concentration was measured with a NanoDrop (Thermo Scientific, Wilmington, USA) and using the extinction coefficient calculated by Protparam[20] from the peptide sequence.

Enzymatic activity assays

The enzymatic activity of recombinant fucosidases was determined on 2'FL, 3FL, DFL, BgA, BgB, BgH, LeA, sLeA, LeX, sLeX, LeY, Fuc1,6GlcNAc, pNP-Fuc and pPGM using 10 µM enzyme, 0.5 mM substrate or 1 mg/mL for pPGM in 50 mM citrate buffer pH 6 and 1 mg/mL bovine serum albumin (BSA). The reactions were incubated for 24 h at 37°C and stopped by boiling at 95°C for 10 min. The release of Fuc was quantified with the k-fucose kit from Megazyme (Wicklow, Ireland) using a microplate reader (FLUOstar Omega, BMG LABTECH, Ortenberg, Germany) by monitoring absorbance at 340 nm every 2 min. To determine the specific activity, the enzymatic reactions were optimised by adjusting enzyme concentration and incubation time to obtain between 6%-25% of substrate hydrolysis and specific activity was calculated from 4 technical replicates. One unit of activity was defined as the amount of enzyme needed to release 1 µmol Fuc per min under the conditions described above. For analysis of reaction products by HPAEC-PAD, the enzymatic reactions were carried out as above but with 0.1 mM substrate, the data were analysed with Prism 5 (GraphPad, San Diego, USA).

For kinetics, all enzymes were incubated with CNP-Fuc in 50 mM citrate buffer at pH 6 and 37°C. Kinetics of TT1386 was also determined with CNP-Fuc in 500 mM glycine buffer at pH 9. The amount of enzyme was determined to fulfil free-ligand approximation, i.e. the enzyme concentration was linear with product formation. The reaction duration was optimised to measure the reaction rates under initial conditions. A standard curve was made with a range of CNP and Fuc from 0 to 0.3 mM. The release of CNP was monitored using a microplate reader (FLUOstar Omega, BMG LABTECH, Ortenberg, Germany) by monitoring absorbance at 405 nm every 2 min for 40 min in 3 technical replicates. The kinetic parameters were calculated based on the Michaelis-Menten equation using a non-linear regression analysis program (Prism 5, GraphPad, San Diego, USA).

Enzymatic reactions (20 µL) were also performed against complex glycans and glycoproteins using 10 µM of enzyme and 5 µM of oligosaccharides or FA2G2 (5 ng/µL)[21], PLA2 (1 mg/mL), IgG (1 mg/mL) untreated or treated with PNGase B035DRAFT_03341 (10 µM)[17] or PNGaseF (5000 units/mL), respectively in 50 mM citrate buffer at pH 6, 37°C for 24 h to release *N*-glycans.

High-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) analysis

GH29 enzymatic reactions were analysed by HPAEC-PAD using a Dionex ICS 5000 system (Thermo Scientific, Hemel Hempstead, UK). The sugars were separated onto a CarboPac PA1 analytical column protected with a CarboPac PA1 guard column using the following gradient conditions: 0 min, 18 mM NaOH; 20–35 min, 100 mM NaOH; 35.1–50 min, 18 mM NaOH.

Liquid chromatography with fluorescence detection and tandem mass spectrometry (LC-FD-MS/MS) analysis

The analysis of enzymatic reactions by -LC-FD-MS/MS was performed as previously described [10]. The reactions were stopped by heating 95°C for 5 min and then dried down using Savant SpeedVac centrifugal evaporator (Thermo Fisher, Wilmington, USA), labelled at the reducing end with procainamide using the glycan labelling kit with sodium cyanoborohydride as the reductant (Ludger, Oxford, UK) and purified using a LudgerClean Procainamide Plate (LC-PROC-96, Ludger, Oxford, UK) to remove the excess dye. The samples were dried down using a Thermo Savant SpeedVac centrifugal evaporator and resuspended in 50 µL of 75% acetonitrile: 25% water. The suspensions were then injected onto a Waters BEH amide column (2.1 x 150 mm, 1.7 µm particle size, 130 Å pore size) at 40 C on a Dionex Ultimate 3000 UHPLC instrument with a fluorescence detector (λex = 310 nm, λem = 370 nm) coupled to a Bruker Amazon Speed ETD. A 50 mM ammonium formate solution pH 4.4 (Ludger, Oxford, UK) was used as mobile phase A and acetonitrile (Romil, UK) was used as mobile phase B. A 70 min gradient was used with mobile phase B from 76–51% from 0 to 53.5 min at a flow rate of 0.4 mL/min followed by mobile phase B from 51–0% from 53.5 min to 55.5 min at flow rate of 0. 2 mL/min, and 2 min stabilisation, mobile phase B from 0–76% from 57.5 min to 59.5 min at a flow rate 0.2 mL/min, and 6 min stabilisation, from 65.5 min to 66.5 min, the flow rate was changed back to 0.4 mL/min and then equilibrated for 3.5 min.

Transfucosylation reactions

For transfucosylation, enzymatic reactions with 1 µM enzyme (1.43 µM for TT1819), 180 mM GlcNAc and 18 mM pNP-Fuc were incubated in 20% (v/v) DMSO for 1 h at 37°C. The reactions were stopped by addition of ethanol using three times the volume of the reaction. To assay the capacity of the enzymes to carry out further transfucosylation reactions, 1 µM enzyme (1.43 µM was used for TT1819) was incubated with 180 mM Fuc1,3GlcNAc or Fuc1,6GlcNAc and 18 mM pNP-Fuc in 20% (v/v) DMSO.

Thin layer chromatography (TLC) and TLC-electrospray ionization-mass spectrometry (TLC-ESI-MS) analysis

To analyse the products of transfucosylation reactions, Fuc (0.01 µmol), pNP-Fuc (0.03 µmol), GlcNAc (0.25 µmol), Fuc1,4GlcNAc (0.005 µmol), Fuc1,3GlcNAc (0.005 µmol), Fuc1,6GlcNAc (0.005 µmol) and the reaction sample (8 x 0.5 µL), were loaded on a 12 cm tall plate (TLC Silica gel 60 F254, Sigma-Aldrich, Germany). The plates were developed using an isopropanol-ammonium hydroxide-water 6:3:1 mixture (namely IPA-NH$_4$OH-H$_2$O) for 3 h or until the frontline of the solvent rose to ca. 11.25 cm. The plate was then dried using a hair

dryer and stained using a 5% ethanolic solution of sulphuric acid. Gently heating of the plate allowed the identification of the TLC spots corresponding to controls and reaction products.

To analyse the products of further transfucosylation reactions, Fuc (0.01 µmol), pNP-Fuc (0.03 µmol), GlcNAc (0.25 µmol), Fuc1,4GlcNAc (0.005 µmol), Fuc1,3GlcNAc (0.005 µmol), Fuc1,6GlcNAc (0.005 µmol), reaction sample with GlcNAc (8 x 0.5 µL), reaction sample with Fuc1,3GlcNAc (4 x 0.5 µL) and reaction sample with Fuc1,6GlcNAc (4 x 0.5 µL) were analysed by TLC as described above.

TLC-ESI-MS of the enzymatic reactions was performed using an Expression Compact Mass Spectrometer (Advion, UK) coupled with a Plate Express reader (Advion, UK) in positive mode to identify the fucosylated reaction products. The enzymatic reactions were analysed through TLC as described above. The analysis was performed in duplicates to stain one TLC plate and use it as a guide to perform the TLC-ESI-MS on the non-stained plate. By comparison with the stained plate, the laser of the Plate Express reader was aimed at the right retention factor (R_f) and the data obtained was analysed using Advion Mass Express software. Comparison of retention factors, in combination with TLC-ESI-MS analysis and NMR allowed identification of reaction products.

Nuclear Magnetic Resonance (NMR) spectroscopy

An aliquot of the enzymatic reactions (600 µL) was evaporated to dryness and reconstituted in 600 µL of NMR buffer (100 mL D_2O containing 0.26 g NaH_2PO_4, 1.41 g K_2HPO_4, and 1 mM deuterated trimethylsilyl propionate (TSP) as a reference compound) before ^{1}H-NMR spectroscopic analysis. ^{1}H-NMR spectra were recorded using a 600-MHz Bruker Avance spectrometer fitted with a 5-mm TCI proton-optimized triple resonance NMR inverse cryoprobe and autosampler (Bruker). Sample temperature was controlled at 300 K. Spectra were acquired with 32 scans, a spectral width of 12500 Hz and an acquisition time of 2.6 s. The "noesypr1d" presaturation sequence was used to suppress the residual water signal with a low-power selective irradiation at the water frequency during the recycle delay. Spectra were then transformed with a 0.3-Hz line broadening and zero filling, manually phased, baseline corrected, and referenced by setting the TSP−d4 signal to 0 ppm. Metabolites were identified by comparison with the spectra of standards (GlcNAc, pNP-Fuc, Fucose, Fuc1,6GlcNAc, Fuc1,3GlcNAc and Fuc1,4GlcNAc).

Saturation transfer difference (STD) NMR

All NMR binding experiments were performed at 278 K on a Bruker Avance III 800 MHz spectrometer equipped with a 5-mm TXI 800 MHz H-C/N-D-05 Z BTO probe. First, FA2G2 was spectroscopically characterised by standard COSY (*cosydfesgpph*), TOCSY (*mlvevphpp*), ^{1}H-^{13}C HSQC (*hsqctgpsp*) and NOESY (*noesygpph*) for the purpose of assignment. Then, FA2G2 was recovered and prepared in a Shigemi advanced NMR microtube assembly at the concentration of ~ 200 µM in the presence of ~ 20 uM TT1819 (protein:ligand ratio 1:20), in D_2O buffer solution containing 25 mM Tris-d_{11} pH 7.8 and 100 mM NaCl. An STD NMR pulse sequence including 2.5 ms and 5 ms trim pulses and a 3 ms spoil gradient was used. Saturation was achieved by applying a train of 50 ms Gaussian pulses (0.40 mW) on the f2 channel, at 6.70 ppm (on-resonance experiments) and 40 ppm (off-resonance experiments). The broad protein signals were removed using a 40 ms spinlock (T1ρ) filter. As a first test for binding, an STD NMR experiment with a saturation time of 2 s and a relaxation delay of 5 s was performed. Then, an STD build up curve was performed, by carrying out STD

experiments at different saturation times (0.5, 1, 2, 3, 4 and 5 s) with 2 K scans, in order to obtain the binding epitope mapping. The resulting build-up curves for each proton were fitted mathematically to a mono-exponential equation ($y = a*[1-\exp(b*x)]$), from which the initial slopes (a*b) were obtained. Finally, the binding epitope mapping was obtained by dividing the initial slopes by the strongest signal corresponding to the methyl group of GlcNAc A, to which an arbitrary value of 100% was assigned.

X-ray crystallography

TT1819 was dialysed into 20 mM Tris 150 mM NaCl. Sitting drop vapour diffusion plates were set up with a protein concentration of 20 mg/mL and 5 mM 2'FL. Crystals appeared in many conditions across commercial sparse matrix screens with the best diffracting crystals appearing in the following condition: 0.12 M diethylene glycol, 0.12 M triethylene glycol, 0.12 M tetraethylene glycol, 0.12 M pentaethylene glycol, 100 mM Tris(base)/bicine pH 8.5, 12.5% 2-methyl-2,4-pentanediol, 12.5% PEG 1000, 12.5% w/v PEG 3350. Diffraction datasets were collected at Diamond Light Source on beamline I24 at a wavelength of 0.9686 Å. Attempts were made to soak out the resulting Fuc molecule bound in the active site so that additional complexes could be attained. These attempts were unsuccessful. By seeding with WT TT1819 crystals we also grew diffracting crystals of TT1819 D218N active site mutant. These crystals were grown in 0.1 M sodium acetate pH 4.6, 8% (w/v) PEG 4000. Data were processed using xia2[22] and dials[23]. The phase problem was solved by molecular replacement using the search model 1ODU, prepared using Chainsaw[24]. Initial model building was performed using ArpWarp[25], followed by alternating cycles of model building and refinement using coot[26], refmac[27], and PDBredo[28]. The refined WT TT1819 structure has 0.3% ramachandran outliers. The final D218N TT1819 structure has 0.05% ramachandran outliers.

Results

Sequence similarity network (SSN) revealed substrate-specificity GH29 fucosidase clusters

SSN was used to explore the sequence-function space of microbial fucosidases belonging to the GH29 family (www.cazy.org). The SSN is composed of nodes and edges, with each representative node representing a single protein sequence, which is linked with an edge when sharing over 40% sequence identity. The SSN analysis of GH29 amino acid sequences revealed a total of 2971 representative nodes wired by 141732 edges. The network was composed of 63 distinct main clusters and 121 singletons defined by cluster analysis utility[29] (**Fig. 1**). Clusters 1, 2, and 3 accounted for 54% of the total nodes. Of the 63 clusters analysed, clusters 1-11, 13, 16, 18, 20, 21, 23, 26, 34, 41, 45 and 47 included sequences corresponding to functionally characterised enzymes. Among them, clusters 1, 13, and 45 contained GH29-B enzymes while the remaining clusters belonged to GH29-A apart from cluster 11, in which Fuc30 isolated from breast-fed infant faecal microbiome was found unrelated to GH29-A or GH29-B subfamilies[14]. Clusters 1 and 13 contained α1,3/4 fucosidases active towards α1,3/4 fucosylated GlcNAc found in Lewis antigens (**Supplementary Table S1**). The convergency ratios of clusters 2, 3 and 4 were lower than 0.30, indicating that these clusters were not isofunctional. Consistent with this, fucosidases belonging to clusters 2, 3, and 4 have been reported to have promiscuous activities for α1,2/3/4/6 fucosyl linkages (**Supplementary Table S1**). Fucosidases in clusters 2 and 3 have been reported to release Fuc from xyloglucans[8,13,30,31]. Cluster 2 also contained the newly found exo-α-l-galactosidase BpGH29 from *Bacteroides plebeius* DSM 17135 [32]. Fucosidases from clusters 3 and 47

as well as non-clustered FucWf4 from *Wenyingzhuangia fucanilytica* CZ1127[T] have been shown to release terminal α1,3/4 Fuc from sulfated fucooligosaccharides[33,34]. Most fucosidases in cluster 4 are of animal origin. Cluster 5 contained fucosidases that specifically act on α1,3 fucosyl linkages with cFase I from *Elizabethkingia meningoseptica* FMS-007 cleaving α1,3 Fuc from the core GlcNAc position from intact glycoproteins [35]. Clusters 6 contained two characterised fucosidases, BF0810 from *Bacteroides fragilis* NCTC 9343 active on pNP-Fuc but not on natural substrates with α1,2/3/4/6 linkages[36]; and Fuc5372 isolated from breast-fed infant faecal microbiome, with preference for α1,2 fucosyl linkages found in HMOs and blood group antigens[14]. Clusters 7, 8 and 10 contained fucosidases with relatively high catalytic efficiency towards aryl-Fuc and marginal activity against α1,2/3/4 fucosyl linkages[7,13,36−38]. Cluster 9 contains Fuc1584 from breast-fed infant faecal microbiome which acts on α1,3/4/6 fucosyl linkages[14]. Clusters 11 and 41 contained α1,6 specific fucosidases with no activity to α1,2/3/4 fucosyl linkages[14,39]. In cluster 16, AlfB from *Lactobacillus casei* BL23 has been reported to be over 800-fold more active on α1,3 fucosylated GlcNAc than on α1,4 fucosylated GlcNAc with the non-terminal Gal in LeX abrogating its activity [39]. Cluster 26 contained site-specific core α1,6 fucosidase AlfC from *L. casei* BL23[39]. Cluster 45 contained Afc1 from *Clostridium perfringens* ATCC 13124 which showed no activity against all aryl- and natural substrates tested[40]. Functionally characterised fucosidases displaying transfucosylation activities were found in GH29-A clusters including clusters 2, 3, 7, 8, 18 and 26 and in cluster 1 belonging to GH29-B subfamily (for full information on functionally characterised fucosidases identified in SSN clusters, see **Supplementary Table S1**). The novel microbial-derived GH29 sequences identified in this SSN analysis belong to a range of microorganisms from Proteobacteria, Actinobacteria, Planctomycetes, Spirochaetota, Firmicutes, Bacteroidetes phyla (**Supplementary Table S2**).

Based on this analysis, we selected 11 GH29 sequences predicted to encode novel fucosidases including three from the GH29-B subfamily, TT1377 and TT1380 in cluster 1, TT4202 in cluster 44; seven from GH29-A subfamily, TT1379 in cluster 2, TT1817 and TT4187 in cluster 3, TT4225 in cluster 4, TT4197 from cluster 9, TT1819 and TT1820 in cluster 26; and one singleton, TT4206. We also included previously characterised fucosidases as controls i.e. two α1,3/4 fucosidases from cluster 1, E1_10125 from *R. gnavus* E1[10] and SsFuc (TT1385) from *Streptomyces* sp. 142; TfFuc1 (TT1386) α1,2/6 fucosidase from *Tannerella forsythia* ATCC 43037 from cluster 8[13,41]; and Afc1 (TT4199) from *C. perfringens* ATCC 13124 from cluster 45, a predicted fucosidase but with no reported activity against any of the α1,2/3/4/6 fucosylated substrates tested[40].

Microbial GH29 enzymes show broad substrate specificity towards fucosylated substrates

The genes encoding the selected GH29 fucosidases were heterologously expressed in *E. coli* and the His$_6$-tag recombinant proteins purified by IMAC and gel filtration (**Supplementary Fig. S1**). *E. coli* Tuner DE3 pLacI strain was chosen as heterologous host as it does not display any endogenous β-galactosidase activity (due to the deletion of the LacZ gene) that may interfere with the enzymatic characterization of the recombinant enzymes.

The kinetic parameters of all GH29 enzymes (TT1377, TT1379, TT1380, TT1385, TT1386, TT1817, TT1819, TT1820, TT4187, TT4199, TT4197, TT4202, TT4206, TT4225 and E1-10125) were determined by calculating the initial rate of reaction with increasing CNP-Fuc concentrations (**Table 1**). All enzymes were found to be

active towards CNP-Fuc, apart from TT4199 from cluster 45 as reported earlier [40]. TT1379, belonging to cluster 2, showed highest activity towards CNP-Fuc among all GH29 enzymes tested, with k_{cat}/K_m of 58.24 $\mu M^{-1} \cdot min^{-1}$, in a range similar to that of Ssα-fuc from the neighbouring node (k_{cat}/K_m=10.25 $\mu M^{-1} \cdot min^{-1}$)[42]. The lowest K_m values were obtained for GH29-A enzymes such as TT1386, TT1817 and TT1820 from clusters 8, 3, 26, respectively, a common feature of GH29-A subfamily enzymes against aryl substrates [4]. Other GH29-A fucosidases were distributed across clusters 3, 4, 8 and 9 have similar K_m values in the range of 153.7 to 668 μM^{-1} whilst k_{cat} values varied from 0.23 to 1411 min^{-1} (**Table 1**). The kinetic parameters of TT1377, TT1380, TT1385 and E1-10125 from cluster 1, and TT4202 from cluster 44 were in the same range with catalytic efficiencies between 10^{-2} and 10^{-1} $\mu M^{-1} \cdot min^{-1}$, consistent with other GH29-B fucosidases i.e. BT1625 from *B. thetaiotaomicron* VPI-5482[9], Eo0918 from *Emticicia oligotrophica* DSM 17448[43], Blon_2336 from *B. longum* subsp. *infantis* ATCC 15697[7].

Next, the substrate specificity of the recombinant fucosidases was tested on a range of fucosylated oligosaccharides. The specific activity was first determined based on fucose release against 2'FL (Fucα1,2Galβ1,4Glc), 3FL (Galβ1-4[Fucα1-3]Glc), DFL (Fucα1-2Galβ1-4[Fucα1-3]Glc), BgA (GalNAcα1-3[Fucα1-2]Galβ1-4GlcNAc), BgB (Galα1-3[Fucα1-2]Galβ1-4GlcNAc), BgH (Fucα1-2Galβ1-4GlcNAc), LeA (Galβ1-3[Fucα1-4]GlcNAc), sLeA (Neu5Acα2-3Galβ1-3[Fucα1-4]GlcNAc), LeX (Galβ1-4[Fucα1-3]GlcNAc), sLeX (Neu5Acα2-3Galβ1-4[Fucα1-3]GlcNAc), LeY (Fucα1-2Galβ1-4[Fucα1-3]GlcNAc), Fuc1,6GlcNAc (Fucα1-6GlcNAc), pPGM and pNP-Fuc using the k-fucose kit (**Table 2**).

TT1377, TT1380, TT1385 and E1-10125 from cluster 1 were found to be over hundred times more active towards α1,3/4 fucosylated linkages than α1,2 fucosylated linkages, while no detectable activity was shown towards Fuc1,6GlcNAc, in line with other characterised GH29-B enzymes from cluster 1 such as BT_2192 [38] (Table S1). In this cluster, only E1-10125 showed similar activity towards both LeX and sLeX [10] (Table S1&2) and pPGM (**Table 2**), consistent with the Lewis epitopes being capped with sialic acids in type III mucin used in this work[44]. TT4199 (i.e. Afc1 from *C. perfringens* ATCC 13124) from cluster 45 belonging to GH29-B showed no activity towards pNP-Fuc, weak activity against Fuc1,6GlcNAc and over thousand times lower activity towards α1,3/4 substrates compared with GH29-B α1,3/4 fucosidases from cluster 1. The non-clustered TT4206 fucosidase showed an enzymatic profile towards α1,3/4/6-linked fucosylated substrates similar to that of TT4199 but weak activity against pNP-Fuc.

The GH29-A enzymes showed preferences towards different α1,2/6 fucosylated linkages. For instance, the highest Fucα1, 2Gal specific activities around 400 U/μmol towards 2'FL and BgH and 0.7 U/μmol towards pPGM were found for TT1379. TT1386 (i.e. TfFuc1 from *T. forsythia* ATCC 43037) from cluster 8 showed the second highest specific activity towards pNP-Fuc, BgH and pPGM. TT1819 and TT1820 from cluster 26, and TT4197 from cluster 9 showed specificity for 6FN, with TT1819 and TT1820 being hundred times more active against pNP-Fuc than TT4197 (**Table 2**). TT1386 showed dual specific activity towards α1,2/6 fucosylated linkages in BgH-II and Fuc1,6GlcNAc, respectively. No detectable activity towards LeX was found for TT1817 and TT4187 from cluster 3. TT4202 only showed specificity towards α1,6 linkage albeit with low activity. TT4225 showed strict specificity for α1,2-fucosylated linkages. None of the enzymes tested in this study showed detectable activity towards blood group A/B type II antigens under the experimental conditions tested, probably due to steric hindrance from the non-terminal GalNAc/Gal.

HPAEC-PAD analyses confirmed the release of Fuc from all GH29 fucosidases tested on their preferred substrates (**Fig. 2 and SupplementaryFig. S2**). In addition, due to the lower detection limit of HPAEC-PAD, it was possible to identify products below 5 µM, which was not possible using the fucose-kit assay. For example, minor Fuc peaks released from BgA/BgB were showed for 10 of the 15 GH29 enzymes tested in this study including TT1377, TT1379, TT1380, TT1385, TT1386, TT1817, TT1820, TT4187, TT4197 and E1-10125. However, no Fuc could be detected by HPAEC for the enzymatic reactions of TT4199 on 2'FL and BgH-II, TT4225 on LeA and sLeA, and TT4187/TT4202 on sLeA, in agreement with their specific activity.

LC-FD-MS/MS was used to investigate whether TT1819, TT4197 and TT4225 fucosidases could act directly on a1,3/6 core fucosylated glycoproteins. Among the different substrates tested, TT4225 showed activity towards IgG glycan and TT1819 was active towards FA2G2 (**Fig. 3**). None of the enzymes tested showed activity with PLA2 or IgG glycoprotein (**SupplementaryFig. S3**).

Microbial GH29-A fucosidases show transfucosylation activity

To test the transfucosylation capacity of the GH29 fucosidases characterised above, the recombinant enzymes were first assayed using GlcNAc as acceptor and pNP-Fuc as donor. The GH29-A fucosidase ATCC_03833 from *R. gnavus* ATCC 29149[10] showing 73.0 % similar to aLfuk1 from *Paenibacillus thiaminolyticus* (both in cluster 3) was used as control as aLfuk1 was previously shown to catalyse the transfer of α-l-fucosyl moiety to different pNP-glycopyranosides with pNP-Fuc as donor[45]. The analysis of the reaction products by TLC confirmed the formation of transfucosylation product by ATCC_03833 and showed that TT1379 from cluster 2, TT1817 from cluster 3, TT1819 and TT1820 from cluster 26 displayed transfucosylation activity (**Fig. 4A and SupplementaryFig. S4B**). These results are in agreement with the SSN analysis showing that GH29 enzymes with reported transglycosylation activity with GlcNAc as acceptor were distributed in clusters 2, 3, 8, and 26 belonging to GH29-A subfamily. In contrast, none of the GH29-B fucosidases tested showed transfucosylation activity using this acceptor-donor pair. Since the R_f values for Fuc1,3GlcNAc, Fuc1,4GlcNAc and Fuc1,6NAc (0.57, 0.52 and 0.55, respectively) on TLC could not discriminate between the products formed, NMR was used to gain further insights into the linkages of the transfucosylation products. The NMR analysis showed that Fuc1,3GlcNAc was the main product generated by TT1379 although traces of Fuc1,4GlcNAc were detected, consistent with TT1379 showing slightly higher hydrolytic activity towards LeX compared to LeA (**Fig. 4B**). ATCC_03833 and TT1817 from cluster 3 produced Fuc1,6GlcNAc and Fuc1,3GlcNAc but not Fuc1,4GlcNAc (**SupplementaryFig. S4C**), in agreement with other cluster 3 enzymes such as AmGH29A from *Akkermansia muciniphila* ATCC BAA-835 with reported activity towards Fuc1,3GlcNAc but not Fuc1,4GlcNAc[46]. Transfucosylation reactions with a1,6 fucosidases TT1819 and TT1820 resulted in the synthesis of Fuc1,6GlcNAc (**SupplementaryFig. S4C**).

TLC-ESI-MS was carried out to further investigate the transfucosylation capacity of TT1819 (**Fig. 4C**). The product of TT1819 enzymatic reaction with GlcNAc was confirmed to be a fucosylated compound (Fuc1,xGlcNAc, found *m/z* 390.1 for [M+Na]$^+$, calcd for $C_{14}H_{25}NO_{10}Na$ 390.1) (**Fig. 4C**). ATCC_03833 reaction with GlcNAc produced a fucosylated product (Fuc1,xGlcNAc, found m/z 390.8 for [M+Na]+, calcd for C14H25NO10Na 390.1) (**SupplementaryFig. S4A**). Further transfucosylation reactions were performed using Fuc1,3GlcNAc or Fuc1,6GlcNAc as acceptors. Both TT1819 and ATCC_03833 produced bifucosylated products with Fuc1,3GlcNAc but not with Fuc1,6GlcNAc (**Fig. 4C** and**SupplementaryFig. S4A**). TT1819 product of the

reaction with Fuc1,3GlcNAc was confirmed to be a product of fucosylation (Fuc1,x[Fuc1,3]GlcNAc, found m/z 537.0 for $[M+Na]^+$, calcd for $C_{20}H_{35}NO_{14}Na$ 536.2) (**Fig. 4C**). ATCC_03833 reaction product with Fuc1,3GlcNAc was confirmed as a fucosylation product (Fuc1,x[Fuc1,3]GlcNAc, found m/z 537.0 for $[M+Na]^+$, calcd for $C_{20}H_{35}NO_{14}Na$ 536.2) (**SupplementaryFig. S4A**). For both enzymatic reactions with Fuc1,6GlcNAc, the only peak produced corresponded to the acceptor (TT1819: Fuc1,6GlcNAc, found m/z 390.8 for $[M+Na]^+$, calcd for $C_{14}H_{25}NO_{10}Na$ 390.1; ATCC_03833: Fuc1,6GlcNAc, found m/z 390.7 for $[M+Na]^+$, calcd for $C_{14}H_{25}NO_{10}Na$ 390.1) (**Fig. 4C** and**SupplementaryFig. S4A**). From this analysis, it is expected that the product of the TT1819 reaction with Fuc1,3GlcNAc and GlcNAc as acceptors is Fuc1,6[Fuc1,3]GlcNAc, in agreement with the substrate specificity of TT1819 for a1,6 linkages (**Fig. 4D**).

Structural basis for TT1819 fucosidase from *Bifidobacterium asteroides* substrate specificity

In addition to its substrate specificity towards α1-6 linkages and transfucosylation activity reported above, LC-FD-MS/MS analyses showed that TT1819 was active against the decasaccharide FA2G2 (**Fig. 2**) while no activity was detected towards IgG glycan or glycoprotein (**SupplementaryFig. S3**). To further explore TT1819 substrate specificity, the crystal structure of the catalytic domain was solved, demonstrating the $(\alpha/\beta)_8$-fold, typical of GH29 enzymes (**Fig. 5A**) and catalytic features conserved with previously solved GH29 enzymes such as AlfC from *Lactobacillus casei* W56 (**Fig. 5B**). Data collection and refinement statistics are detailed in **Table 3**. It was only possible to grow diffracting crystals of TT1819 in the presence of 2'FL, resulting in a complex with Fuc bound in the active site (**Fig. 5A and SupplementaryFig. S5A**). Asp218 and Asp260 were identified as catalytic nucleophile and acid/base, respectively based on proximity to the Fuc residue and homology with other GH29 enzymes, such as AlfC from *L. casei* W56[47]. Asp218 is flanked by the structurally conserved Tyr151 that donates a hydrogen bond to the nucleophile, as previously observed in TmαFuc and E1_10125 fucosidases from *T. maritima* and *R. gnavus* E1, respectively[10,48]. Extensive hydrogen bonding interactions were observed between the active site and the bound sugar hydroxyl groups. The C6 methyl group sits in a hydrophobic pocket formed by Trp216 and Trp305. Unlike E1_10125, which showed evidence of β-fucose bound[10], the electron density of the TT1819 complex most clearly matched α-fucose. Furthermore, attempting to model β-fucose led to a steric clash with Asp210. High B-factors were observed in the residues surrounding the active site, indicating that there may be plasticity in the presence of larger substrate molecules (**SupplementaryFig. S5B**). However, minimal conformation changes are observed when comparing the Fuc bound WT active site to that of an unbound D218N catalytic mutant (**Fig. 5A**). Compared to fucosidases E1_10125 from *R. gnavus* E1 and Blon_2336 from *B. longum* subsp. *infantis,* the TT1819 active site was shown to be constricted (**SupplementaryFig. S5C, S5D**), which may contribute to the substrate specificity of this enzyme.

Tyr57 is of interest in relation to TT1819 α1,6 linkage specificity. The residue hydrogen bonds with the catalytic acid/base at the active site boundary (**Fig. 5B**). This residue is structurally conserved in AlfC (as Tyr37), found in the same SSN cluster as TT1819 and shows specificity to α1,6-linked Fuc[47]. AlfC Tyr37 has been shown to change conformation in presence of the α1,6-linked ligand. Tyr37 then forms an aromatic subsite, providing a stacking interaction with the monosaccharide that is immediately linked to the intimately held Fuc[47]. This conformational change and function are expected to be maintained by TT1819. In contrast, an equivalent residue is absent in E1_10125 and Blon_2336 from *B. longum* subspecies *infantis*, both belonging to cluster 1.

This cluster favors hydrolysis of α1,3/4 fucosyl linkages rather than α1,6 (**Supplementary Fig. S5C, S5D**). Additionally, it is proposed that TT1819 Ile284 would clash with substrates presenting α1,3/4 linkages, whereas Blon_2336 has an acidic residue well placed to create a stabilising hydrogen bond to the substrate (**Supplementary Fig. S5D**).

To gain further structural insights into the ligand specificity of TT1819, saturation transfer difference nuclear magnetic resonance spectroscopy (STD NMR) studies (Mayer and Meyer, 1999) were conducted with the inactive TT1819 D218A mutant in the presence of FA2G2 (**Fig. 6A**). The D218A mutation allowed the NMR study to focus on the process of molecular recognition of the substrate, disentangling it from the subsequent chemical reaction. Transfer of magnetization as saturation from the protein to the ligand was observed, in agreement with the activity of TT1819 for this substrate. Due to the large size of FA2G2 (decasaccharide), the 1D NMR spectrum showed significant chemical shift overlapping, challenging the analysis. For that reason, only isolated protons were assigned and quantitatively analysed (i.e. protons H5 and H6 of fucose, H2s of mannose and the methyl group of the four GlcNAc rings). A full build-up curve analysis of their STD intensities showed that the enzyme intimately recognises the non-reducing end sugar residues constituting FA2G2 (**Fig. 6B** and **Supplementary Fig. S6**) with no significant differences in their binding epitopes. The main contacts were restricted to Fuc and GlcNAc (**Fig. 6A**) residues, whereas only loose contacts were observed with the distant GlcN moieties (**Fig. 6B, 6C, 6D**).

Discussion

Reflecting the high diversity of fucosylated structures in nature, microbes produce a range of α-l-fucosidases with different linkages specificity[4]. Here, we report the enzymatic characterisation of new GH29 α-L fucosidases identified through SSN analysis, expanding the GH29 enzyme toolbox. GH29 enzymes are divided into GH29-A and GH29-B, displaying broad and narrow substrate specificity, respectively. The acid/base residues of GH29-B enzymes are conserved and assignable from primary sequence alignments[9] in contrast to GH29-A where catalytic residues are less conserved[31,45,47−50]. Here, we showed using SSN that GH29-A sequences were spread within 18 clusters as compared to 3 clusters for GH29-B sequences, in line with the high variability characteristic of GH29-A enzymes. Notably, cluster 1 with 560 fucosidase sequences belonging to GH29-B, and clusters 2 and 3 belonging to GH29-A accounted for over half of the GH29 sequences. However, we observed high level of variation in the catalytic efficiency of GH29-A and GH29-B enzymes towards pNP-Fuc and CNP-Fuc substrates which, although highest with GH29-A fucosidases, is challenging the dogma that GH29-B enzymes are not active on these substrates, as also supported by a recent study exploring functional diversity of GH29 family[15]. Collectively, our results showed that GH29-B enzymes have a preference for Lewis antigen epitopes while the linkage preference of GH29-A enzymes varies between clusters, but fucosidases from both families can display strict linkage preferences.

The recombinant GH29 enzymes characterised in this work spanning different SSN clusters showed substrate specificities in line with functionally characterised GH29 fucosidases from these clusters. For example, TT1377, TT1380 and TT1385 belonging to cluster 1 showed Fuc1,3/4GlcNAc and Fuc1,3Glc linkage preferences albeit with different catalytic efficiency. None of the fucosidases from cluster 1 GH29-B were active on Fuc1,6GlcNAc. TT1379 belonging to cluster 2 showed preference for α1,2Gal linkages and displayed

transfucosylation activity, as also reported for FgFCO1 from *Fusarium oxysporum* 0685 belonging to the same cluster[13,31,51,52]. TT1817 and TT4187 belonging to cluster 3 showed highest activity towards pNP-Fuc but no significant activity was found towards α1,2/3/4/6 fucosylated substrates, as also reported for Fp251, Fp239 and Fp231 from *Paraglaciecola sp.*[53], Alf1_Wf from *W. fucaniytia* CZ1127[T34] and ATCC_03833 from *R. gnavus* ATCC 29149[10] found in the same cluster. TT4225 belonging to cluster 4 showed preference for Fucα1,2Gal, consistent with Fucosidase O from *Omnitrophica* bacterium OLB16[54], HsFucA1[55] and HsFucA2[56] from *Homo sapiens* found in the same cluster. TT4197 belonging to cluster 9 was highly active on Fuc1,6GlcNAc with marginal activity towards pNP-Fuc, similar to Fuc1584 isolated from breast-fed infant faecal microbiome in the same cluster[14]. TT1819 and TT1820 belonging to cluster 26 were active on both Fuc1,6GlcNAc and pNP-Fuc, which might be associated with their transfucosylation activity, as shown with AlfC from *L. casei*[11,12]. TT4202 from cluster 44 (which does not contain any functionally enzymes) and non-clustered TT4206, showed low activity against all substrates tested. Collectively, these data confirmed that SSN is a reliable approach to predict the substrate specificity of GH29 enzymes belonging to characterised clusters. During the preparation of this manuscript, a bioinformatic analysis based on Conserved Unique Peptide Patterns (CUPP) was applied to predict the substrate specificity and transglycosylation capacity of GH29 enzymes[15,57]. Among the novel GH29 α- l- fucosidases characterised as part of their work, BT3665 and BT3956 from *B. thetaiotaomicron* VPI-5482 showed substrate preference for 2'FL and Fuc1,6GlcNAc, respectively[14,15], which is in line with their presence in SSN clusters 2 and 11 respectively, encompassing enzymes of similar substrate specificity. Cluster 11 also includes Fuc30 from breast-fed infant faecal microbiome with the same substrate specificity[14]. WfFuc from *Wenyingzhuangia fucanilytica* from cluster 20 showed activity towards 3FL[15]. Together, these recent data further reinforce the suitability of SSN to predict substrate specificity.

GH29 α-l-fucosidases have a retaining double-displacement mechanism with retention of anomeric configuration[48], allowing the catalysis of transglycosylation reactions leading to the synthesis of oligosaccharides, such as fucosylated HMOs. Previously characterised α-l-fucosidases AlfB and AlfC from *Lacticaseibacillus casei* W56 have been shown to synthesise Fuc1,3GlcNAc, Fuc1,6GlcNAc, the glycoamino acid Fuc1,6GlcNAc, and several 6'-fucosyl-glycans[11,12]. Fucosyl-N-GlcNAc disaccharides have also been produced using *Bacteroides fragilis* α-l-fucosidase[36]. The HMOs, 2'FL, 3FL, and lacto-N-fucopentaose II, have been synthesised in low amounts using α-l-fucosidases from *T. maritima*, *C. perfringens*, and a soil-derived metagenome library[13,58]. Here, we showed that, using pNP-Fuc as donor and GlcNAc as acceptor, TT1819 and TT1820 from cluster 26 produced Fuc1,6GlcNAc as sole transfucosylation product, TT1379 from cluster 2 mainly produced Fuc1,3GlcNAc while TT1817 and ATCC_03833 from produced Fuc1,6GlcNAc and Fuc1,3GlcNAc. The linkage specificity observed during catalysis was retained during transfucosylation. Bi-fucosylated-GlcNAc products were produced by TT1819 with Fuc1,3GlcNAc as donor. These newly characterised GH29 α-l-fucosidases might therefore be exploited as biotechnological tools in the synthesis of oligosaccharides. In addition, our work highlighted the suitability of SSN as a tool to predict transfucosylation capacity since functionally characterised with transglycosylation capacity fall within discrete clusters. This was further supported by findings from the recent CUPP study[15] where newly characterised GH29 α-l-fucosidases with transglycosylation activity are distributed in SSN clusters predicted to include such activity i.e. cluster 2 (BT3665, FgFCO1, NixE from *Xanthomonas campestris* pv. *campestris* str. ATCC 33913), cluster 3

(Mfuc5 from soil metagenome), cluster 8 (TfFuc1), cluster 1 (BbAfcB from *B. bifidum* ATCC 1254 and CpAfc2 from *Clostridium perfringens* ATCC 13124).

Bifidobacteria are common gut commensal bacteria specialised in HMO degradation and metabolism[59]. TT1819 is derived from *Bifidobacterium asteroides*, an ancestor of the genus Bifidobacterium. Although many α-l-fucosidases from Bifidobacteria including *B. bifidum* and *Bifidobacterium longum* subsp. *infantis*[7,60,61] have been identified by bioinformatics analysis, as belonging to GH29, GH95, and GH151 families, few have been functionally characterised. In the GH29 family, bifidobacterial α-l-fucosidases have been divided into GH29-BifA fucosidases (only found in *B. bifidum* strains), GH29-BifB fucosidases, GH29-BifC fucosidases, and GH29-BifD fucosidases based on their domain conversation and phylogeny[62]. BbAfcB belonging to cluster 1, the only characterised representative GH29 fucosidase of GH29-BifA, was active on 3FL, Lewis group antigens (A, B, X, and Y), and lacto-N-fucopentaose II and III but not on glycoconjugates containing α1,2-fucosyl residue or on synthetic pNP-Fuc[61]. The only characterised bifidobacterial GH29-BifB fucosidase (Blon_2336 from *B. longum* subsp. *infantis* ATCC 15697) belonging to cluster 1 revealed similar activity to BbAfcB (GH29-BifA) against Fuc1,3, Fuc1,3GlcNAc, and Fuc1,4GlcNAc linkages[7]. These GH29-BifB fucosidases appear to be distributed across bifidobacterial strains of different species (unlike GH29-BifA fucosidases) and frequently, strains that exhibit GH29-BifB fucosidases also produce GH29-BifC fucosidases. GH29-BifC fucosidases can catalyse the hydrolysis of core α1,6-fucose on the N-glycan of glycoprotein and Fuc1-6GlcNAc-IgG[63]. Among the *B. asteroides* strains which have been genome-sequenced to date, fucosidase-encoding genes are restricted to GH29 family, with one GH29-encoding gene in the DSM 20089 and PRL2011 strains and 3 in the ESL0447 strain (ww.cazy.org). Based on phylogeny analysis, these would fall into GH29-BifD. Here, we showed that TT1819 from *B. asteroides*, in cluster 26, exhibited α1,6 substrate specificity and could hydrolyse FA2G2 *N*-glycan through recognition of Fuc1,6GlcNAc epitopes, as shown by STD NMR. TT1819 crystal structure displays the conserved GH29 catalytic machinery. Comparisons with AlfC from *L. casei* BL23, found in the same SSN cluster, indicated that TT1819 specificity may be due to the presence of the Tyr57-containing loop. Additionally, compared to Blon_2336, the constricted active site in TT1819 through the intrusion of Ile284 will obstruct access to α1,3 and α1,4 linked fucosylated substrates. Mutations of these residues may help modulate enzymatic activity in favour of transfucosylation over hydrolysis, opening a new route towards the synthesis of fucosylated conjugates that may be used as prebiotics for promoting the growth of Bifidobacteria in the gut.

The continuing expansion of microbial GH families within the CAZy database, through metagenomic sequencing, with many uncharacterised or "hypothetical" proteins is an opportunity to identify novel enzymes with biotechnological applications. This work demonstrated the suitability of SSN as a powerful bioinformatics tool to harness the wealth of sequencing data and help predict novel fucosidases and transfucosylation activities in prokaryotes. The combination of GH grouping into CAZy families and SSN clusters provides a strong prediction on their putative biological functions.

Declarations

Data availability statement

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials.

Disclosure statement

No potential competing interest was reported by the authors.

References

1. Henrissat, B. & Davies, G. Structural and sequence-based classification of glycoside hydrolases. Curr. Opin. Struct. Biol. 7, 637–644 (1997).

2. Garron, M. L. & Henrissat, B. The continuing expansion of CAZymes and their families. Curr. Opin. Chem. Biol. 53, 82–87 (2019).

3. Drula, E. et al. The carbohydrate-active enzyme database: Functions and literature. Nucleic Acids Res. 50, D571–D577 (2022).

4. Wu, H., Owen, C. D. & Juge, N. Structure and function of microbial α-l-fucosidases: a mini review. Essays Biochem. EBC2022015, (2023).

5. Ndeh, D. et al. Complex pectin metabolism by gut bacteria reveals novel catalytic functions. Nature 544, 65–70 (2017).

6. Kovaľová, T. et al. The first structure–function study of GH151 α-l-fucosidase uncovers new oligomerization pattern, active site complementation, and selective substrate specificity. FEBS J. 289, 4998–5020 (2022).

7. Sela, D. A. et al. *Bifidobacterium longum* subsp. *infantis* ATCC 15697 α-fucosidases are active on fucosylated human milk oligosaccharides. Appl. Environ. Microbiol. 78, 795–803 (2012).

8. Lezyk, M. et al. Novel α-l-fucosidases from a soil metagenome for production of fucosylated human milk oligosaccharides. PLoS One 11, 1–18 (2016).

9. Shaikh, F. A., Lammerts Van Bueren, A., Davies, G. J. & Withers, S. G. Identifying the catalytic acid/base in GH29 α-l-fucosidase subfamilies. Biochemistry 52, 5857–5864 (2013).

10. Wu, H. et al. Fucosidases from the human gut symbiont *Ruminococcus gnavus*. Cell. Mol. Life Sci. 78, 675–693 (2021).

11. Rodríguez-Díaz, J., Carbajo, R. J., Pineda-Lucena, A., Monedero, V. & Yebra, M. J. Synthesis of fucosyl-*N*-Acetylglucosamine disaccharides by transfucosylation using α-l-Fucosidases from *Lactobacillus casei*. Appl. Environ. Microbiol. 79, 3847–3850 (2013).

12. Becerra, J. E. et al. Unique microbial catabolic pathway for the human core *N*-glycan constituent fucosyl-α-1,6-*N*-acetylglucosamine-asparagine. MBio 11, 1–18 (2020).

13. Zeuner, B. et al. Substrate specificity and transfucosylation activity of GH29 α-l-fucosidases for enzymatic production of human milk oligosaccharides. N. Biotechnol. 41, 34–45 (2018).

14. Moya-Gonzálvez, E. M. et al. Infant gut microbial metagenome mining of α-l-fucosidases with activity on fucosylated human milk oligosaccharides and glycoconjugates. Microbiol. Spectr. 10, (2022).

15. Perna, V. N., Barrett, K., Meyer, A. S. & Zeuner, B. Substrate specificity and transglycosylation capacity of α-l-fucosidases across GH29 assessed by bioinformatics-assisted selection of functional diversity. Glycobiology 1–15 (2023). doi:10.1093/glycob/cwad029

16. Grootaert, H., van Landuyt, L., Hulpiau, P. & Callewaert, N. Functional exploration of the GH29 fucosidase family. Glycobiology 00, 1–11 (2020).

17. Crouch, L. I. et al. Plant *N*-glycan breakdown by human gut Bacteroides. Proc. Natl. Acad. Sci. U. S. A. 119, 1–11 (2022).

18. Huang, Y., Niu, B., Gao, Y., Fu, L. & Li, W. CD-HIT Suite: A web server for clustering and comparing biological sequences. Bioinformatics 26, 680–682 (2010).

19. Gerlt, J. A. et al. Enzyme function initiative-enzyme similarity tool (EFI-EST): A web tool for generating protein sequence similarity networks. Biochim. Biophys. Acta - Proteins Proteomics 1854, 1019–1037 (2015).

20. Gasteiger, E. et al. Protein Identification and Analysis Tools on the ExPASy Server. in The Proteomics Protocols Handbook 571–607 (Humana Press, 2005). doi:10.1385/1-59259-890-0:571

21. Rohrer, J. S., Basumallick, L. & Hurum, D. C. Profiling N-linked oligosaccharides from IgG by high-performance anion-exchange chromatography with pulsed amperometric detection. Glycobiology 26, 582–591 (2016).

22. Winter, G., Lobley, C. M. C. & Prince, S. M. Decision making in xia2. Acta Crystallogr. Sect. D Biol. Crystallogr. 69, 1260–1273 (2013).

23. Winter, G. et al. DIALS: Implementation and evaluation of a new integration package. Acta Crystallogr. Sect. D Struct. Biol. 74, 85–97 (2018).

24. Stein, N. CHAINSAW: A program for mutating pdb files used as templates in molecular replacement. J. Appl. Crystallogr. 41, 641–643 (2008).

25. Cohen, S. X. et al. ARP/wARP and molecular replacement: The next generation. Acta Crystallogr. Sect. D Biol. Crystallogr. 64, 49–60 (2007).

26. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 486–501 (2010).

27. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 355–367 (2011).

28. Joosten, R. P., Joosten, K., Murshudov, G. N. & Perrakis, A. PDB-REDO: Constructive validation, more than just looking for errors. Acta Crystallogr. Sect. D Biol. Crystallogr. 68, 484–496 (2012).

29. Zallot, R., Oberg, N. & Gerlt, J. A. The EFI web resource for genomic enzymology tools: leveraging protein, genome, and metagenome databases to discover novel enzymes and metabolic pathways. Biochemistry 58, 4169–4182 (2019).

30. Curci, N. et al. Xyloglucan oligosaccharides hydrolysis by exo-acting glycoside hydrolases from hyperthermophilic microorganism *Saccharolobus solfataricus*. Int. J. Mol. Sci. 22, (2021).

31. Cao, H., Walton, J. D., Brumm, P. & Phillips, G. N. Structure and substrate specificity of a eukaryotic fucosidase from *Fusarium graminearum*. J. Biol. Chem. 289, 25624–25638 (2014).

32. Robb, C. S. et al. Metabolism of a hybrid algal galactan by members of the human gut microbiome. Nat. Chem. Biol. 18, 501–510 (2022).

33. Silchenko, A. S. et al. Fucoidan-active α-l-fucosidases of the GH29 and GH95 families from a fucoidan degrading cluster of the marine bacterium *Wenyingzhuangia fucanilytica*. Arch. Biochem. Biophys. 728, 109373 (2022).

34. Dong, S., Chang, Y., Shen, J., Xue, C. & Chen, F. Purification, expression and characterization of a novel α-l-fucosidase from a marine bacteria *Wenyingzhuangia fucanilytica*. Protein Expr. Purif. 129, 9–17 (2017).

35. Li, T. et al. Identification and characterization of a core fucosidase from the bacterium *Elizabethkingia meningoseptica*. J. Biol. Chem. 293, 1243–1258 (2018).

36. Liu, P. et al. Screening and characterization of an α-l-fucosidase from *Bacteroides fragilis* NCTC9343 for synthesis of fucosyl-*N*-acetylglucosamine disaccharides. Appl. Microbiol. Biotechnol. 104, 7827–7840 (2020).

37. Shi, R. et al. Biochemical characterization of a novel α-l-fucosidase from *Pedobacter* sp. and its application in synthesis of 3'-fucosyllactose and 2'-fucosyllactose. Appl. Microbiol. Biotechnol. 104, 5813–5826 (2020).

38. Sakurama, H. et al. Differences in the substrate specificities and active-site structures of two α-l-fucosidases (Glycoside hydrolase family 29) from *Bacteroides thetaiotaomicron*. Biosci. Biotechnol. Biochem. 76, 1022–1024 (2012).

39. Rodríguez-Díaz, J., Monedero, V. & Yebra, M. J. Utilization of natural fucosylated oligosaccharides by three novel α-l-fucosidases from a probiotic *lactobacillus casei* strain. Appl. Environ. Microbiol. 77, 703–705 (2011).

40. Fan, S. et al. Cloning, characterization, and production of three α-l-fucosidases from *Clostridium perfringens* ATCC 13124. J. Basic Microbiol. 56, 347–357 (2016).

41. Megson, Z. A. et al. Characterization of an α-l-fucosidase from the periodontal pathogen *Tannerella forsythia*. Virulence 6, 282–292 (2015).

42. Cobucci-Ponzano, B., Trincone, A., Giordano, A., Rossi, M. & Moracci, M. Identification of an archaeal α-l-fucosidase encoded by an interrupted gene: Production of a functional enzyme by mutations mimicking programmed – 1 frameshifting. J. Biol. Chem. 278, 14622–14631 (2003).

43. Liu, S. et al. The fucosidase-pool of *Emticicia oligotrophica*: Biochemical characterization and transfucosylation potential. Glycobiology 26, 871–879 (2016).

44. Glenister, D. A., Salamon, K. E., Smith, K., Beighton, D. & Keevil, C. W. Enhanced growth of complex communities of dental plaque bacteria in mucin-limited continuous culture. Microb. Ecol. Health Dis. 1, 31–38 (1988).

45. Kovaľová, T., Bene, E., Lipovová, P. & Dohnálek, J. Active site complementation and hexameric arrangement in the GH family 29; a structure – function study of α-l-fucosidase isoenzyme 1 from *Paenibacillus thiaminolyticus*. Glycobiology 29, 59–73 (2019).

46. Shuoker, B. et al. Sialidases and fucosidases of *Akkermansia muciniphila* are crucial for growth on mucin and nutrient sharing with mucus-associated gut bacteria. Nat. Commun. 14, 1833 (2023).

47. Klontz, E. H. et al. Structure and dynamics of an α-fucosidase reveal a mechanism for highly efficient IgG transfucosylation. Nat. Commun. 11, 1–14 (2020).

48. Sulzenbacher, G. et al. Crystal structure of *Thermotoga maritima* α-l-fucosidase: Insights into the catalytic mechanism and the molecular basis for fucosidosis. J. Biol. Chem. 279, 13119–13128 (2004).

49. Bueren, A. L. Van et al. Analysis of the reaction coordinate of alpha-l-fucosidases: a combined structural and quantum mechanical approach. J. Am. Chem. Soc. 132, 1804–6 (2010).

50. Armstrong, Z., Meek, R. W., Wu, L., Blaza, J. N. & Davies, G. J. Cryo-EM structures of human fucosidase FucA1 reveal insight into substrate recognition and catalysis. Structure 30, 1443–1451.e5 (2022).

51. Schopohl, D. et al. Purification and properties of a secreted and developmentally regulated α-l-fucosidase from *Dictyostelium discoideum*. J. Biol. Chem. 267, 2400–2405 (1992).

52. Biel-Nielsen, T. L. et al. Utilization of industrial citrus pectin side streams for enzymatic production of human milk oligosaccharides. Carbohydr. Res. 519, 108627 (2022).

53. Schultz-Johansen, M., Stougaard, P., Svensson, B. & Teze, D. Characterization of five marine family 29 glycoside hydrolases reveals an α-l-fucosidase targeting specifically Fuc(α1,4)GlcNAc. Glycobiology 32, 529–539 (2022).

54. Vainauskas, S. et al. A novel broad specificity fucosidase capable of core α1–6 fucose release from *N*-glycans labeled with urea-linked fluorescent dyes. Sci. Rep. 8, 9504 (2018).

55. Dawson, G. & Tsay, G. Substrate specificity of human α-l-fucosidase. Arch. Biochem. Biophys. 184, 12–23 (1977).

56. Dicioccio, R. A., Barlow, J. J. & Khushi, L. Substrate specificity and other properties of a-l-fucosidase from human serum. J. Biol. Chem. 257, 714–718 (1982).

57. Barrett, K., Hunt, C. J., Lange, L., Grigoriev, I. V & Meyer, A. S. Conserved unique peptide patterns (CUPP) online platform 2.0: implementation of + 1000 JGI fungal genomes. Nucleic Acids Res. 1–7 (2023). doi:10.1093/nar/gkad385

58. Zeuner, B. & Meyer, A. S. Enzymatic transfucosylation for synthesis of human milk oligosaccharides. Carbohydr. Res. 493, 108029 (2020).

59. Thongaram, T., Hoeflinger, J. L., Chow, J. M. & Miller, M. J. Human milk oligosaccharide consumption by probiotic and human-associated Bifidobacteria and Lactobacilli. J. Dairy Sci. 100, 7825–7833 (2017).

60. Turroni, F. et al. Genome analysis of *Bifidobacterium bifidum* PRL2010 reveals metabolic pathways for host-derived glycan foraging. Proc. Natl. Acad. Sci. U. S. A. 107, 19514–19519 (2010).

61. Ashida, H. et al. Two distinct α-l-fucosidases from *Bifidobacterium bifidum* are essential for the utilization of fucosylated milk oligosaccharides and glycoconjugates. Glycobiology 19, 1010–1017 (2009).

62. Curiel, J. A. et al. Architecture insight of bifidobacterial α-l-fucosidases. Int. J. Mol. Sci. 22, 1–15 (2021).

63. Ashida, H. et al. 1,6-α-l-fucosidases from *Bifidobacterium longum* subsp. *infantis* ATCC 15697 involved in the degradation of core-fucosylated *N*-glycan. J. Appl. Glycosci. 67, 23–29 (2020).

Tables

Table 2 is available in the Supplementary Files section

Table 1 Kinetic parameters of GH29 fucosidases on CNP-Fuc.

GH29 Michaelis-Menten parameters

| GH29	sub-family	cluster	Vmax (μM/min)			Km (μM)			kcat (min⁻¹)	kcat/Km (μM⁻¹·min⁻¹)
TT1377	B	1	2.25	±	2.26E-02	302.10	±	9.69	112.60	0.37
TT1379	A	2	3.52	±	2.57E-02	302.30	±	7.05	17605.00	58.24
TT1380	B	1	2.60	±	4.06E-02	289.50	±	14.61	51.90	0.18
TT1385	B	1	4.78	±	0.23	341.60	±	51.67	4.78	1.40E-02
TT1386 pH6	A	8	0.53	±	1.98E-02	153.70	±	21.81	534.70	3.48
TT1386 pH9	A	8	1.41	±	7.34E-02	185.60	±	34.88	1411.00	7.60
TT1817	A	3	4.47	±	0.20	208.10	±	33.24	894.60	4.30
TT1819	A	26	1.45	±	3.83E-02	222.00	±	20.10	726.50	3.27
TT1820	A	26	1.07	±	5.17E-02	188.10	±	32.87	533.00	2.83
TT4187	A	3	12.13	±	0.46	668.00	±	69.04	12.13	1.82E-02
TT4199	B	45		-			-		-	-
TT4197	A	9	14.47	±	1.74	495.20	±	171.80	144.70	0.29
TT4202	B	44	4.80	±	0.13	299.60	±	25.34	47.96	0.16
TT4206	*	n.c.	2.19	±	4.63E-02	290.60	±	19.81	0.44	1.51E-03
TT4225	A	4	0.56	±	6.71E-03	279.80	±	10.82	0.23	8.06E-04
E1_10125	B	1	1.93	±	3.29E-02	248.20	±	12.27	20.04	7.38E-02

*, unknown

n.c., non-clustered

3 technical replicates were performed for each GH29-CNP-Fuc reaction.

Table 3 X-ray crystallography data collection and refinement statistics

|	D218N Apo	WT Fucose bound
PDB identifier	8P1S	8P1R
Data collection		
Space group	P21212	P21212
Cell dimensions		
a, b, c (Å)	87.7, 136.3, 160.5	89.1, 142.8, 167.2
α, β, g (	90, 90, 90	90, 90, 90
Resolution (Å)	103.86-1.66 (1.69-1.66)	108.56-1.90 (1.93-1.90)
R_{merge}	0.20 (0.8)	0.25 (9.9)
I/sigI	11.2 (1.9)	9.2 (0.4)
Completeness %	99.1 (94.5)	100 (100)
Redundancy	21.1 (7.8)	17.5 (14.9)
CC(½)	0.997 (0.541)	0.998 (0.215)
Refinement		
Resolution	104.07-1.66	83.74-1.90
No. reflections	224428	167720
R_{work}/R_{free}	0.193/0.160	0.201/0.176
No. atoms		
Protein	16417	16209
Ligand	40	122
Water	1942	749
B-factors		
Protein	16.8	27.6
Ligand	18.9	44.9
Waters	23.5	38.36
r.m.s deviations		
Bond lengths (Å)	0.014	0.008
Bond angles (°)	1.8	1.33

Figures
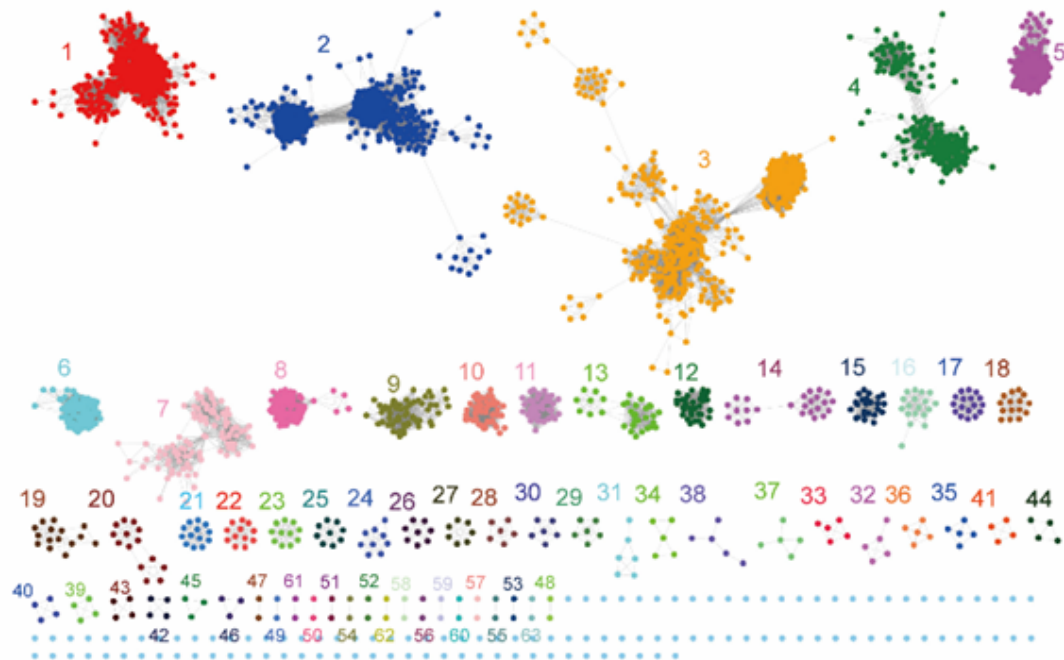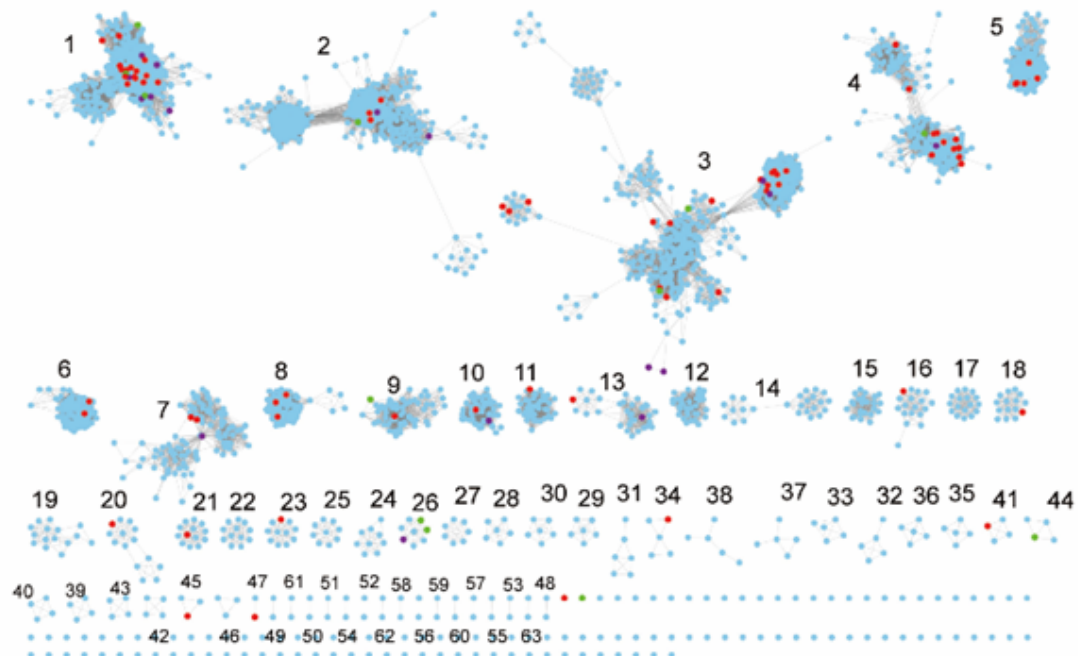
Figure 1

Sequence similarity network GH29 fucosidases family. A) The coloured SSN of GH29 family after cluster analysis. B) The distribution of functionally characterised GH29s in different clusters. Red nodes represent enzymatically characterised GH29s, purple nodes represent structurally characterised GH29s while green nodes represent new GH29 enzymes characterised in this work.
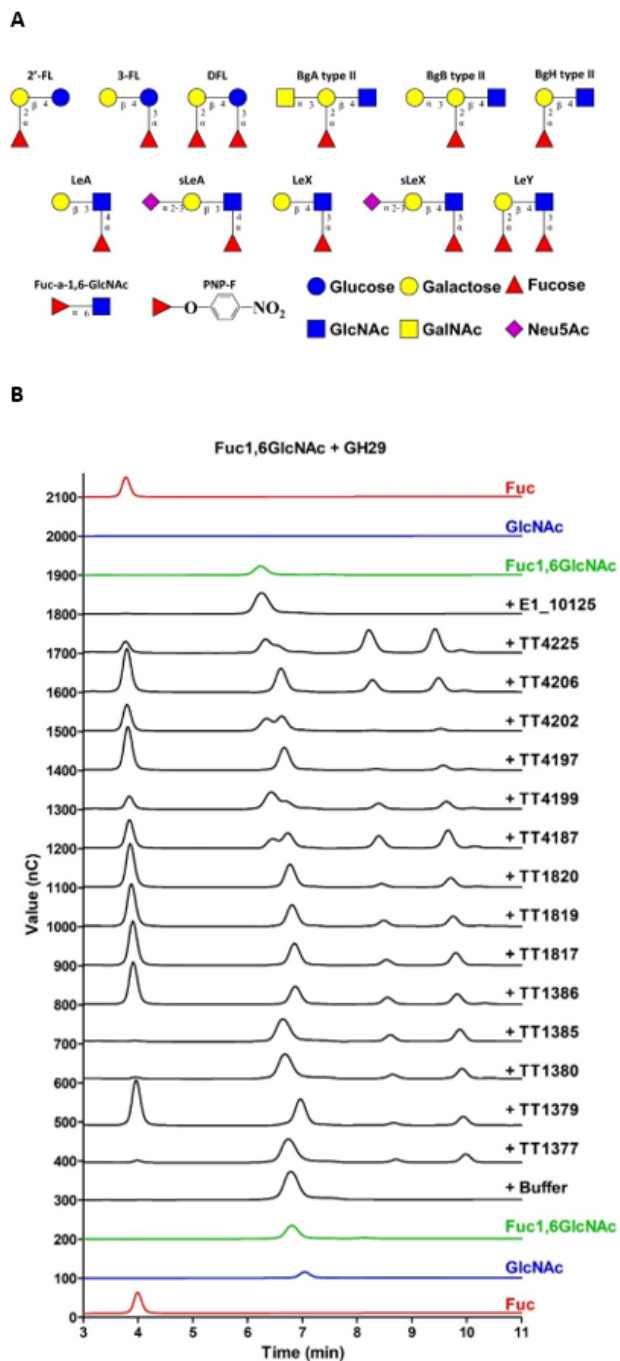
Figure 2

Analysis of GH29 fucosidase reactions on fucosylated substrates. A) Fucosylated oligosaccharides used in this study. Monosaccharide symbols follow the Symbol Nomenclature for Glycans system. **B**) HPAEC-PAD analysis of GH29 enzymatic reaction with Fuc1,6GlcNAc. The data were analysed with Prism. Standards were Fuc (red), Fuc1,6GlcNAc (green), GlcNAc (blue). The black lines correspond to the enzymatic reactions with Fuc1,6GlcNAc incubated with the different GH29 fucosidases tested or in the presence of buffer. For HPAEC enzymatic reaction analyses of GH29 enzymes with all other substrates depicted in A), see Supplementary Figure S2.
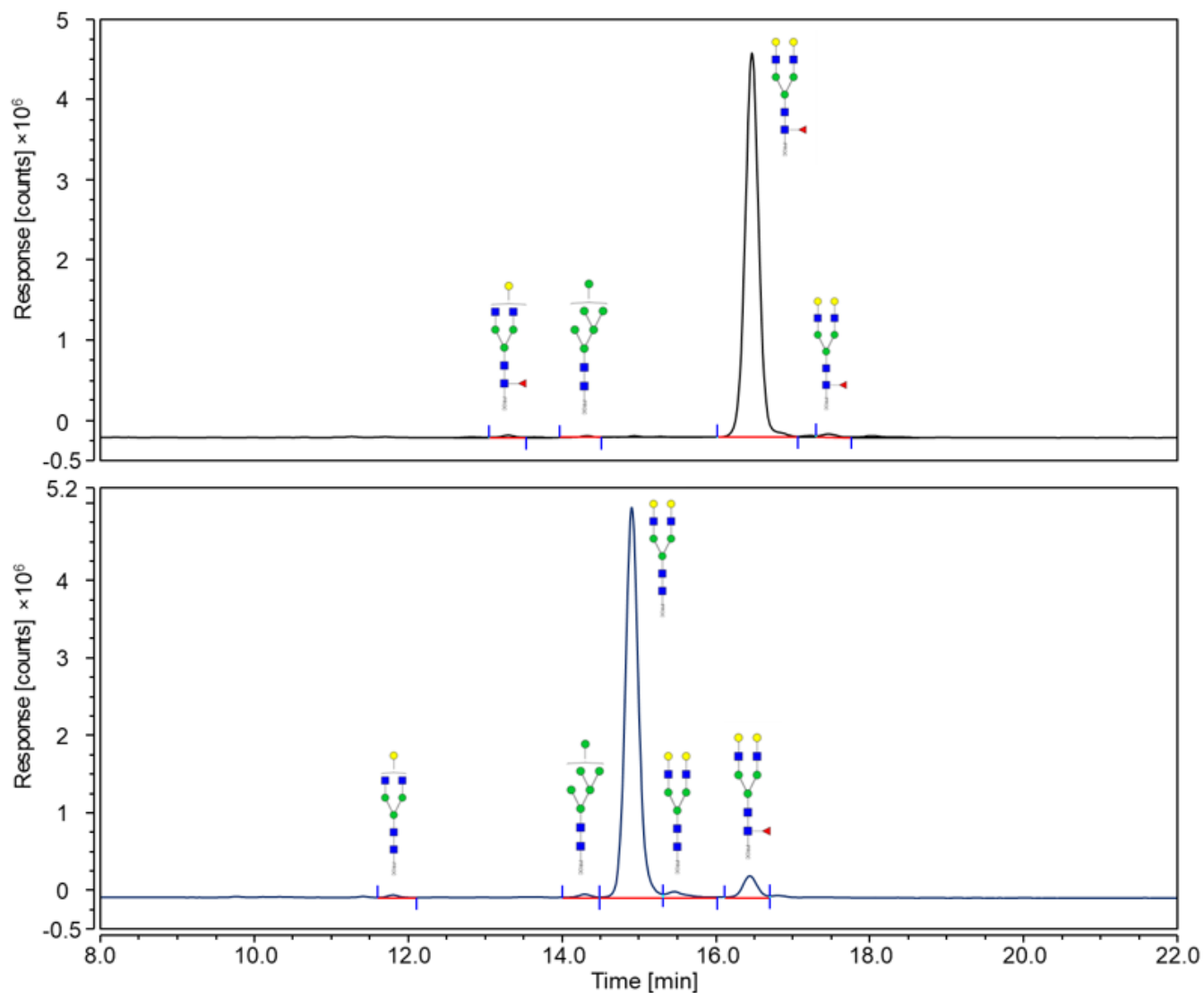
Figure 3

LC-FD-MS/MS analysis of TT1819 reaction with FA2G2. The upper and lower panels correspond to reactions with and without enzyme, respectively. Glycan products are annotated next to peaks on the chromatograms.
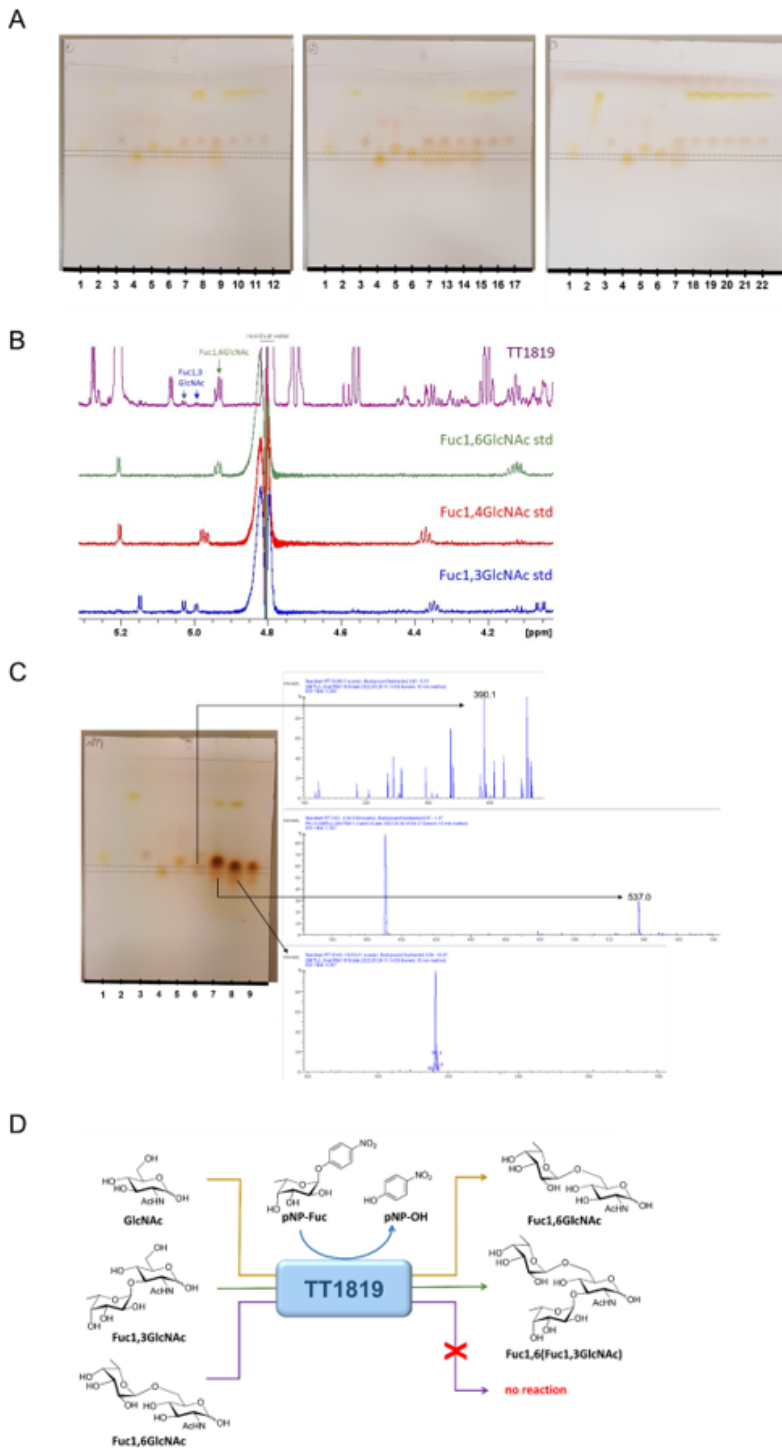
Figure 4

Transfucosylation activity of GH29 fucosidases. **A**) TLC analysis of GH29 transfucosylation reactions with GlcNAc as acceptor and pNP-Fuc as donor; ATCC_03833 was used as control. Lanes 1 to 6 correspond to standards: Fuc (lane 1), pNP-Fuc (lane 2), GlcNAc (lane 3), Fuc1,4GlcNAc (lane 4), Fuc1,3GlcNAc (lane 5) and 6 is Fuc1,6GlcNAc (lane 6). Lane 7 is the control reaction with ATCC_03833. Lanes 8 to 22 are the GH29 reactions, TT1377 (lane 8), TT1379 (lane 9), TT1380 (lane 10), TT1385 (lane 11), TT1386 (lane 12), TT1817 (lane 13), TT1819 (lane 14), TT1820 (lane 15), TT4187 lane 16), TT4199 (lane 17), TT4197 (lane 18), TT4202 (lane 19), TT4206 (lane 20), TT4225 (lane 21) and E1_10125 (lane 22). **B**) 600MHz ^{1}H NMR spectra of

TT1819 reaction and standards of Fuc1,3GlcNAc, Fuc1,4GlcNAc and Fuc1,6GlcNAc. The mid field region displays distinctive signals showing the presence of Fuc1,6GlcNAc and trace levels of Fuc1,3GlcNAc in TT1819. **C**) TLC and TLC-ESI-MS analysis of TT1819 transfucosylation reactions with GlcNAc, Fuc1,3GlcNAc or Fuc1,6GlcNAc as acceptors and pNP-Fuc as donor. Lanes 1 to 5 and 9 are standards, Fuc (lane 1), pNP-Fuc (lane 2), GlcNAc (lane 3), 4 is Fuc1,4GlcNAc (lane 4), Fuc1,3GlcNAc (lane 5) and Fuc1,6GlcNAc (lane 9). Lanes 6 to 8 are TT1819 reactions, with GlcNAc (lane 6), Fuc1,3GlcNAc (lane 7) and Fuc1,6GlcNAc (lane 8). **D**) Schematic of the TT1819 enzymatic fucosylation reaction using GlcNAc, Fuc1,3GlcNAc and Fuc1,6GlcNAc as acceptors and pNP-Fuc as donor.
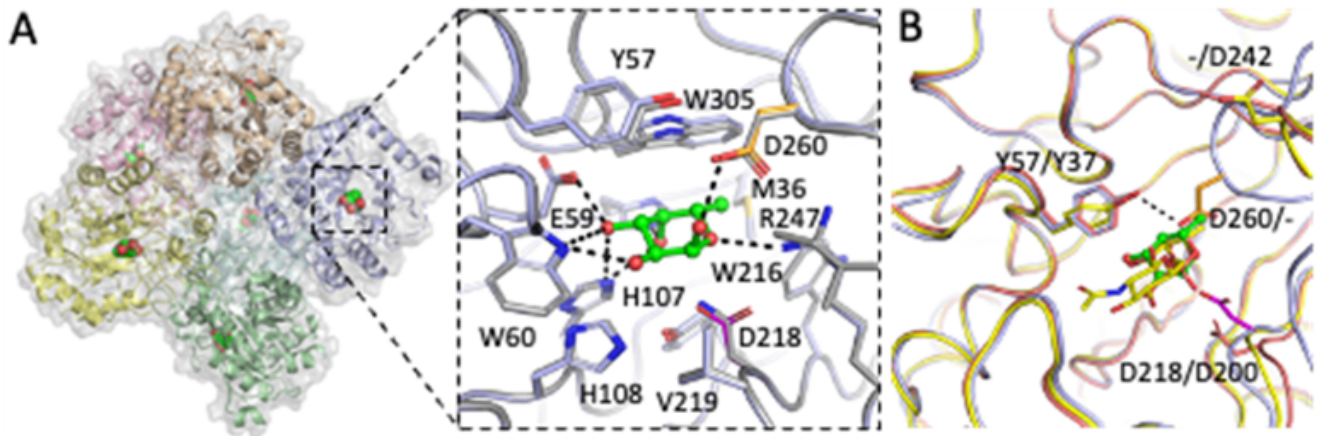

Figure 5

Crystal structure of TT1819. **A**) Crystal structure of TT1819 in complex with Fuc. Boxout shows ligand bound WT TT1819 in light blue and unbound D218N in grey. The bound Fuc residue is shown in green. The catalytic acid base and nucleophile residues are highlighted in orange and magenta, respectively. Hydrogen bonding interactions are indicated with black dashed lines. **B**) Proposed rotation of active site Tyr57 in the presence of substrate molecules suggested by alignment to AlfC, show bound to Fuc1,6GlcNAc in yellow (PDB 6OHE) and bound to fucose in pink (PDB 6O1A). TT1819 is shown in light blue.
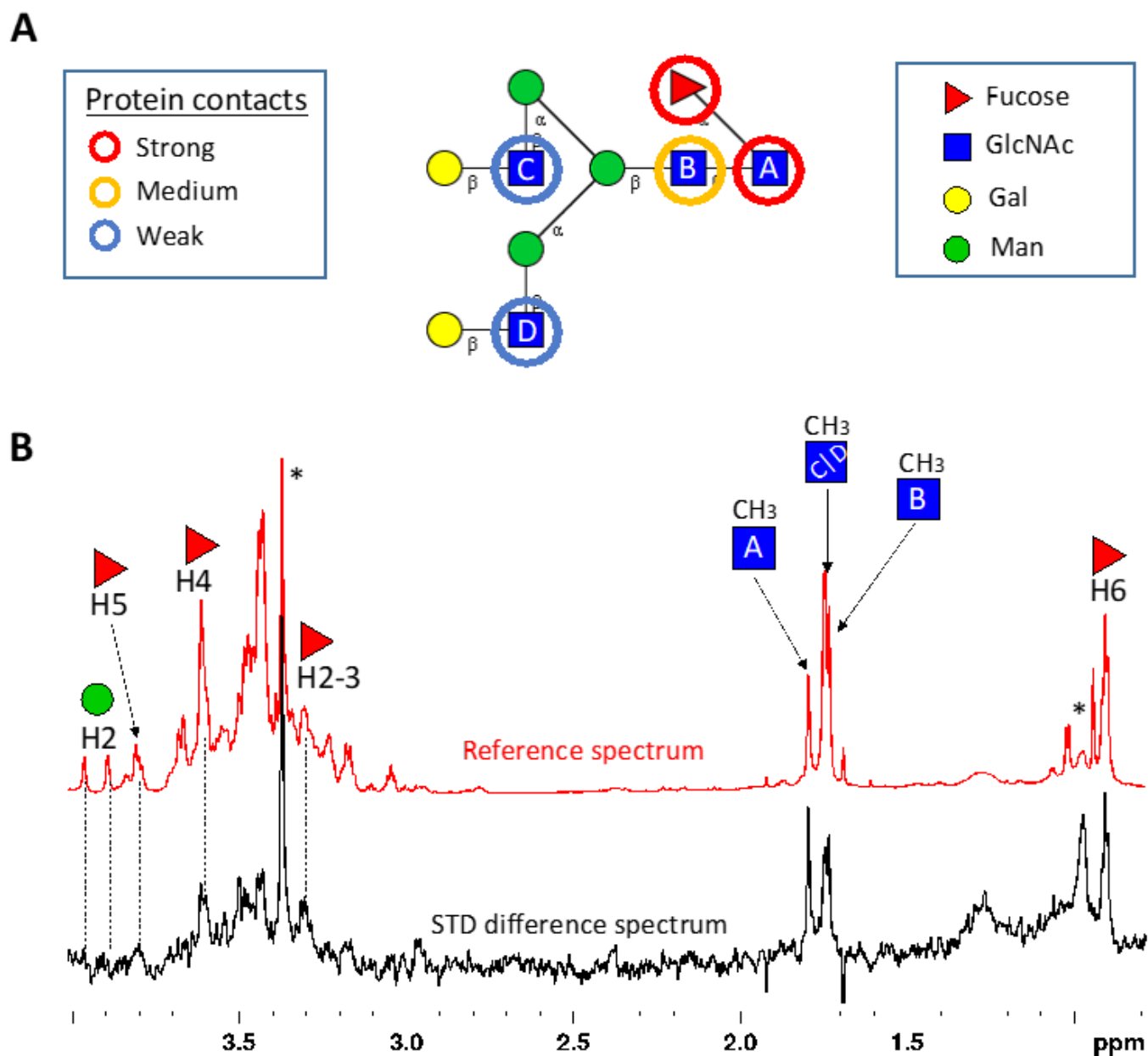
Figure 6

STD-NMR analysis of the interaction between TT1819 and FA2G2. **A**) Binding epitope mapping of FA2G2 as bound to TT1819 from STD NMR experiments. Protein contact strength reflects relative values of saturation transfer after normalization to the most intense one (the methyl group of GlcNAc(A)) obtained from STD initial slopes (full STD NMR build-up curves and initial slopes for each proton can be found in the **Supplementary Data**). **B**) STD NMR difference (black) and reference (red) spectra of the FA2G2/TT1819 D218A sample, acquired at 2 s saturation time. Only isolated protons that were unambiguously assigned could be analysed for binding epitope determination and are labelled on the spectra (impurities are marked with *). The STD NMR analysis supports that the enzyme preferentially recognizes the reducing end, with sugar rings of Fuc and GlcNAc(A) showing the strongest STD intensities.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- GH29SupplementaryFinal.docx
- Table2.docx