

Exploring the dynamic role of L-fucose and α -L-fucosidases in the cross talk between gut bacteria and host

Yvette Magdalena Cornelia Adriana Luijkx

Copyright © 2021 **Y.M.C.A. Luijckx**

All rights reserved. No part of this thesis may be reproduced, stored or transmitted in any way or by any means without the prior permission of the author, or when applicable, of the publishers of the scientific papers.

Layout and design | Eduard Boxem | www.persoonlijkproefschrift.nl

Printing: Ridderprint | www.ridderprint.nl

Exploring the dynamic role of L-fucose and α -L-fucosidases in the cross talk between gut bacteria and host

Onderzoeken van de dynamische rol van L-fucose en α -L-fucosidases in de interactie tussen darm bacteriën en de gastheer

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

maandag 5 juli 2021 des middags te 12.15 uur

door

Yvette Magdalena Cornelia Adriana Luijkx

geboren op 21 december 1991
te Etten-Leur

Promotor:

Prof. dr. G.J.P.H. Boons

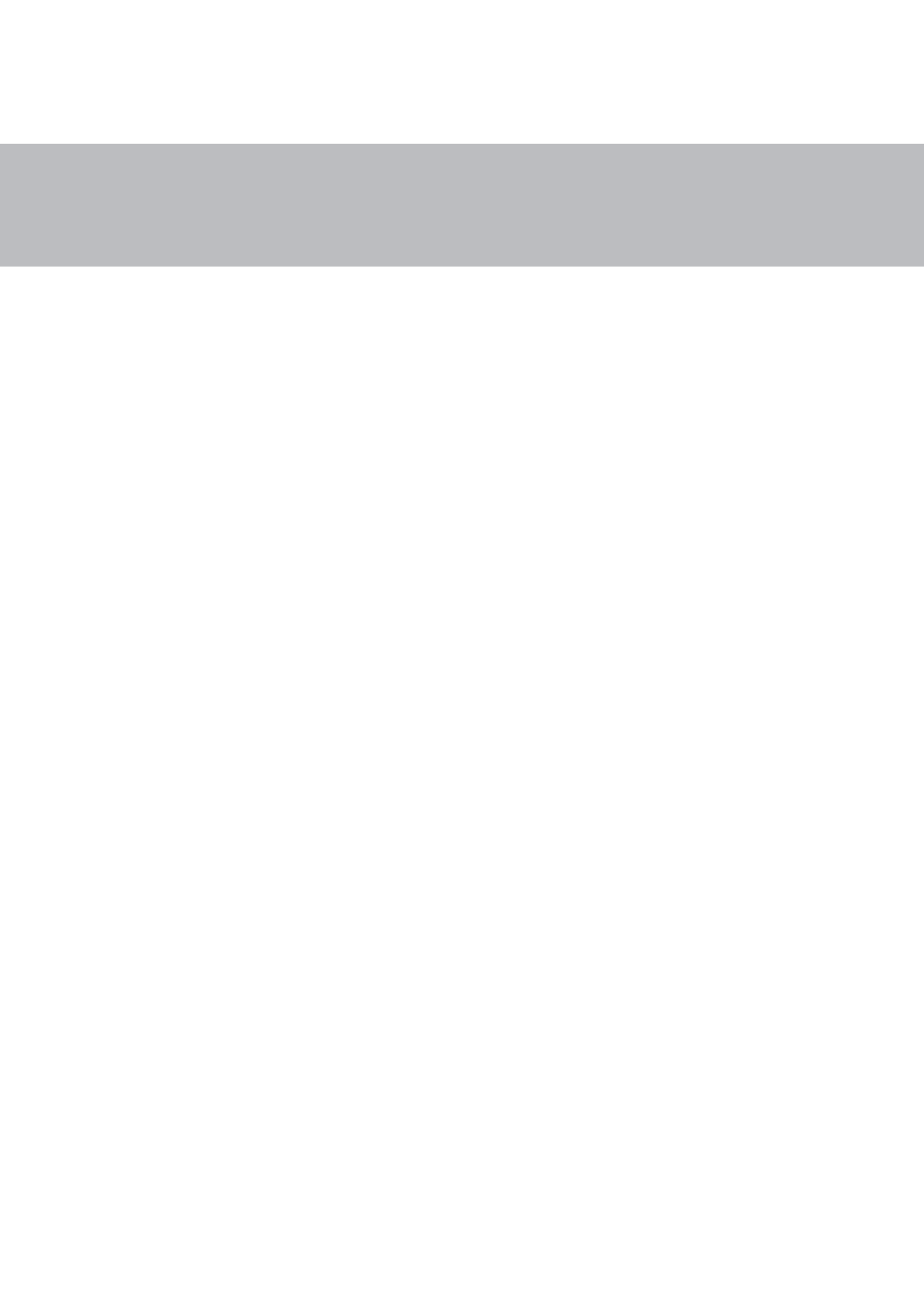
Copromotoren:

Dr. T. Wennekes

Dr. K. Strijbis

Table of Contents

Chapter 1	General introduction	7
Chapter 2	<i>Bacteroides fragilis</i> fucosidases facilitate growth and invasion of <i>Campylobacter jejuni</i> in the presence of mucins	25
Chapter 3	O-glycomic signatures of sodium butyrate-stimulated enterocyte differentiation	45
Chapter 4	Development of a 1,2-difluorofucoside activity-based probe for profiling GH29 fucosidases	69
Chapter 5	Detection of bacterial α -L-fucosidases with an <i>ortho</i> -quinone methide-based probe and mapping of the probe-protein adducts	91
Chapter 6	Summary, Future Perspectives and Concluding Remarks	177
Appendix	Nederlandse Samenvatting	144
	Curriculum Vitae	149
	List of Publications	150
	Acknowledgement	151



Chapter 1

General introduction

The human gut epithelial cells express glycans, especially $\alpha 1,2$ -fucosyl linkages, which function as a biological interface for the human-gut microbiota interactions. The research described in this thesis was conducted with the aim of developing and implementing activity-based probes (ABPs) that target fucosidase enzymes at the interface of the intestinal epithelial cells and surrounding microbiota. Fucosidases are glycoside hydrolases that specifically catalyze the removal of L-fucose. Fucosylated glycoproteins are abundant at the gut-microbiota interface and are increasingly being investigated because of their involvement in critical biological processes such as immune response, signal transduction, and adhesion of pathogens. Understanding the role of L-fucose in the interactions between our gut and microbiota requires quantification and visualization of the enzymes involved in altering this fucosylation pattern. The first part of this chapter provides an introduction and detailed description of the composition of the human gut microbiota interface. This is followed by a brief overview of the current literature on fucosidases expression in relation to health and disease. Next, the composition of gut microbiota and relevant studies investigating the activity of their fucosidases are discussed. Following this, the catalytic mechanisms used by the different classes of fucosidases will be discussed and how these insights are translated to the designs of currently known ABPs. The chapter ends with an outline for the following chapters.

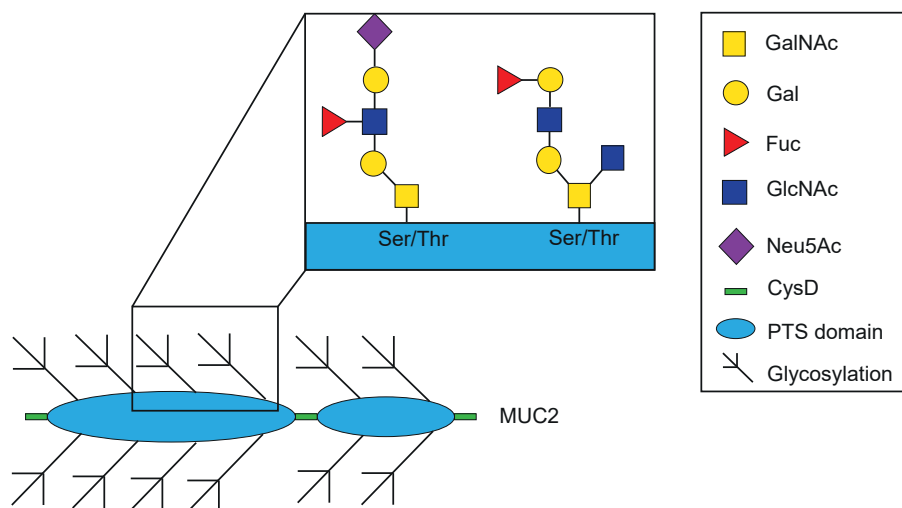


Figure 1 | Schematic representation of the MUC2 protein. The protein consist of two domains rich in serine-, threonine-, proline- (PTS domain) that serves as attachment sites for *O*-linked glycans. These PTS domains are interrupted by short cysteine rich domains (CysD).

1.The human gut – microbiota interface

The human gastrointestinal (GI) tract forms a specific niche for a complex and dynamic population of microbes that contribute significantly to the maintenance of health and the onset and progression of disease.¹ The human GI tract consist of epithelial cells that are covered by a mucus layer. This mucus layer shields the epithelial cells from both the healthy microbiota that are continuously present in the lumen as well as incoming enteric pathogens. A bacterial metabolite known to be a regulator of mucus production, is butyrate.^{2,3} Butyrate is the product of bacterial degradation of undigested complex carbohydrates, and forms the primary energy source of colonocytes.⁴ Mucus is predominantly composed of gel-forming mucins, which are complex *O*-glycosylated proteins secreted by goblet cells within the epithelial lining.⁵ Mucin genes are differently expressed throughout the body and 50% of the mass of mucin proteins consists of *O*-glycosylated sugar chains. In the colon, the most abundant secreted mucin is MUC2, which is composed of 80% *O*-glycans (Figure 1). Besides the mucin composition also the microbiota density and diversity differ along the GI tract.^{6,7} The highest amount of microbiota can be found in the large intestine, particularly the colon. In the colon, the mucus layer is divided into two distinct layers. The inner mucus layer, close to the epithelium is tightly crosslinked and virtually free of bacteria. The outer mucin layer, on the other hand, is more open in structure and provides a nutrient-rich niche for the microbiota (Figure 2).⁸

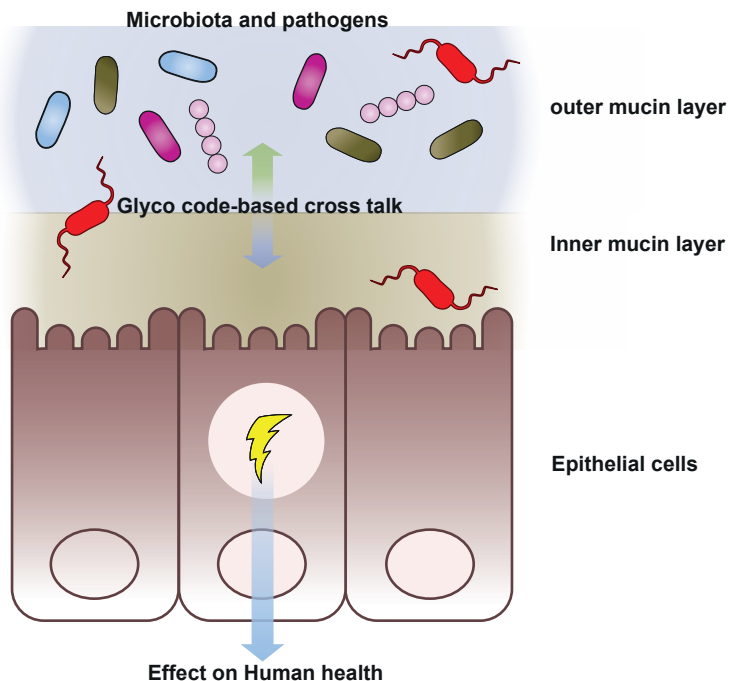


Figure 2 | Schematic overview of the cross talk between microbes and the host gut epithelial cells.

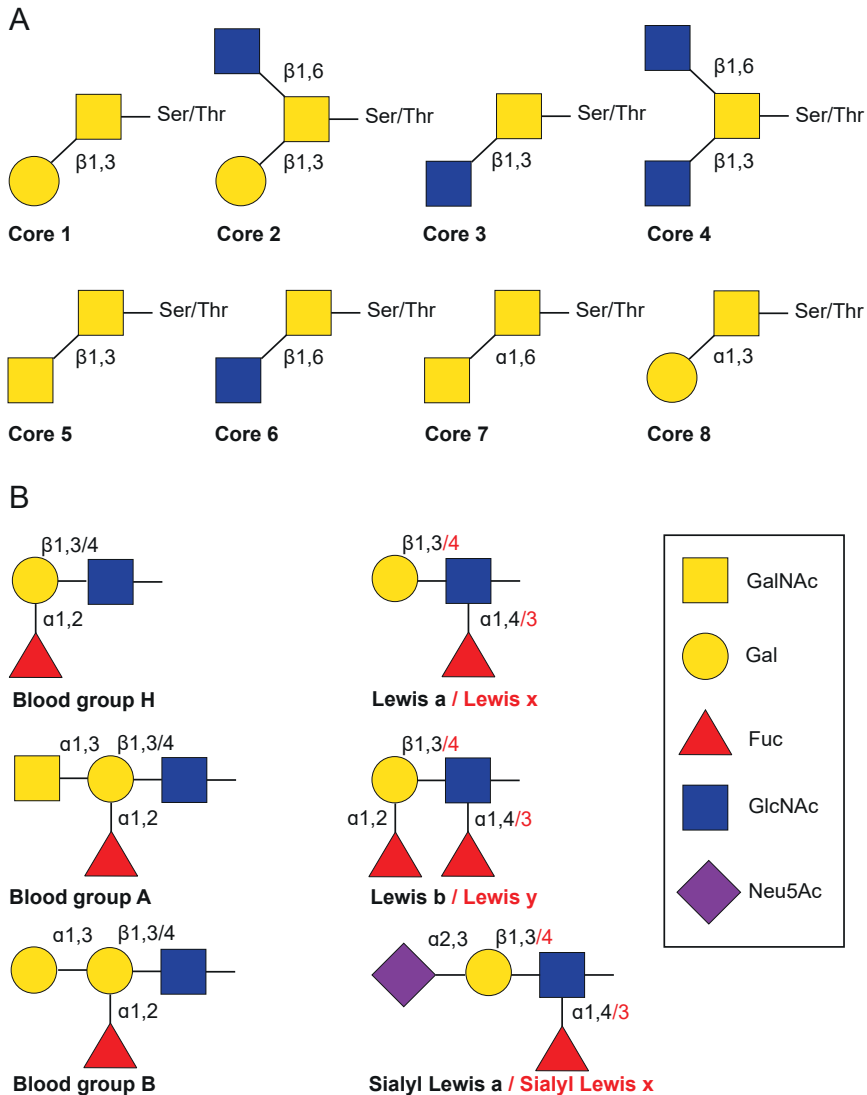


Figure 3 | Schematic representation of GI mucin glycans. (a) The common mucin-type *O*-glycans (core 1-8) found in the GI tract. (b) Main glycan epitopes in GI mucins.

1.1 Composition intestinal mucins and glycosylation

The family of mucin glycoproteins consists of the members MUC1 to MUC21, which are expressed differentially along the GI tract and at other mucosal surfaces.^{9,10} One mucin protein can harbour more than 100 different *O*-glycan structures.^{11–13} The synthesis of mucin oligosaccharides starts with the transfer of *N*-acetylgalactosamine (GalNAc) to

Ser and Thr residues on the mucin protein. Eight core structures of the mucin *O*-linked glycans have been identified, with cores 1-4 most commonly found on intestinal mucins (Figure 3A).¹⁴ Human gastric mucins generally contain core 1 and core 2 structures, while core 3 structures are predominant in the small intestine, and core 3 and 4 together make up the majority of structures found on the colonic mucins. These core structures can be extended with *N*-acetylglucosamine (GlcNAc), galactose (Gal), *N*-acetylgalactosamine (GalNAc), fucose (Fuc) or sialic acid (Neu5Ac) monosaccharides with the latter two occupying the majority of terminal positions (Figure 3B). These terminal epitopes are the main source of glycan diversity present on the human gut and form a specific glycosylation pattern (Figure 3). Noticeable is that the ratio of sialic acid to fucose in humans increases along the GI tract, from the ileum to the rectum, and an inverse gradient is seen in mice.^{15,16}

2. Fucosidases in health and disease

It is well established that the mucosa glycosylation, formed by the high glycan diversity, is critical in the maintenance of a homeostatic relationship between the gut microbiota and their host.^{17–21} Intestinal bacteria are known to scavenge host glycans (e.g., carbohydrate structures from *O*-glycosylated mucins and other human blood group antigens (HBGA) on mucosal surfaces).²² Therefore, the availability of these glycan structures plays an important role in biological processes, such as assisting commensal colonization.^{23,24} Furthermore, 20% of the world wide population is affected by a so called “non-secretor” phenotype which enables them to synthesize H-type antigens, the precursor of the HBGA.²⁵ The non-secretor status is caused by nonsense polymorphisms of the FUT2 gene, which mediates the addition of L-fucose via an α 1,2 linkage to the terminal β -D-galactose residues of mucosal glycans, causing defects in addition of α 1,2-fucose. Intestinal microbiome analysis revealed that non-secretors showed an aberrant gut microbiota compared to secretors.²⁶ Additionally, it has been shown that non-secretors are more susceptible to infection by pathogenic microbes.^{25,27–29} Besides the FUT2 secretor status several other factors influencing glycosylation have shown to play important roles in the maintenance of a homeostatic relationship.

In particular, bacterial α -L-fucosidases are key enzymes for the modification of intestinal mucin glycans by gut microbes.²² This enzymatic activity allows access to L-fucose as nutrient, which contributes to the adaption of gut bacteria to specific niches (Figure 4B). This was for example shown for the human commensal *Ruminococcus gnavus* which expresses a fucosidase with a specificity for sialic acid terminated fucosylated glycans (Lewis X/A). This specificity may have contributed to the adaption of this commensal to the infant gut.³⁰ Besides enabling colonization of gut commensals, abnormal activity of α -L-fucosidase enzymes has been associated with several pathological conditions, such as inflammation, tumor progression and bacterial infections.^{31–36} It has previously been reported that human α -L-fucosidase is released in response to *Helicobacter pylori* infection. Interestingly, the response of the pathogen is to use the released L-fucose as

an additional energy source and for the production of host like Lewis X structures on its own surface to escape immune surveillance.³⁴ Cross feeding between the fucosidase-expressing epithelial cells or commensal bacteria and invading pathogens, seems to be a common strategy that enhances the virulence of various pathogenic bacteria (Figure 4A). The common foodborne pathogen *Salmonella typhimurium* upregulates its fucose-catabolizing genes quickly after infection, indicating that it is metabolizing fucose.³⁷ As it lacks fucosidases, it cannot access fucose from the complex glycan structures, but must rely on other bacteria that secrete fucosidases. Competition experiments in mice infected with either wildtype *S. typhimurium* or a mutant strain with a deletion in the fucose-utilization pathway showed a significant expansion disadvantage for the mutant.³⁷ Other studies including our own have shown that also for the well-known pathogen *Campylobacter jejuni*, L-fucose metabolism, enabled by cross feeding, seems to be important for survival in the gut.^{38–40}

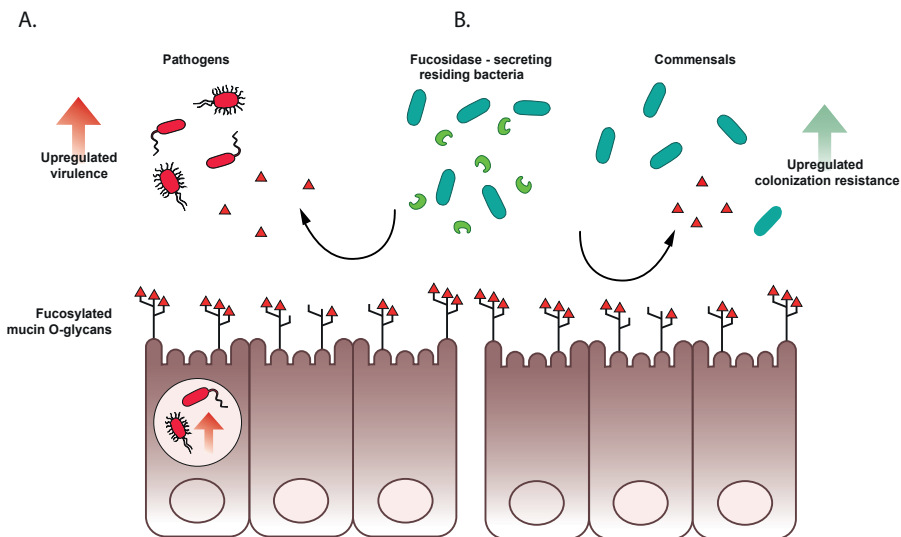


Figure 4 | Schematic model of secreted fucosidases enhancing growth and invasion of both pathogens and commensals. Fucosylated mucin O-glycans (red triangles) are expressed on the apical surface of intestinal epithelial cells. Residing bacteria such as *Bacteroides thetaiotamicron* secrete fucosidases (green), thereby liberating L-fucose from the mucin O-glycans. (a) It was hypothesized that pathogens such as *H. pylori* and *S. typhimurium* scavenge and metabolize L-fucose, liberated by others, which increases their virulence. (b) Liberation of L-fucose leads to an increased colonization of the commensal species belonging to the Bacteroidetes phyla.

2.3 Fucosidase expression of intestinal microbiota

For a long period, mucin degradation was thought to be associated with pathogenicity. However, in the past decade research has shown the presence and importance of mucin degradation by intestinal commensals.⁴¹ The healthy human gut is heavily colonized by bacteria with most species belonging to the phyla *Bacteroidetes*, *Actinobacteria*,

Verrucomicrobia, and *Firmicutes*.⁴² Here I outlined the studies into some of these main intestinal species in correlation to their mucin degradation properties, with particular attention to fucosidases activity.

Mucin degradation by intestinal bacteria has been extensively studied in the Bacteroidetes phyla. Early studies on *Bacteroides thetaiotaomicron* showed that all strains tested were able to ferment glycosaminoglycans.⁴³ Later it was shown that a specific *B. thetaiotaomicron* strain was also capable of growing on glycans purified from Porcine Gastric Mucins (PGM).⁴⁴ Transcriptomics analysis highlighted genes coding for putative glycoside hydrolases (GHs) such as α -L-fucosidase, endo- β -N-acetylglucosaminidase, endo- β -galactosidase and α -mannosidase upregulated in the presence of PGM.^{44,45} This showed its ability to adapt to liberating and utilizing host sugars, including L-fucose. Interestingly, it has been shown that *B. thetaiotaomicron* upregulates its fucosidases activity and fucose metabolism genes when living in mice compared to growth *in vitro*.⁴⁶ Furthermore, the bacterium also stimulates host fucosylation, likely to be used as an energy source to sustain its own colonization.⁴⁷ Another species in the Bacteroidetes phyla, *Bacteroides fragilis*, is capable of growing on mucins as a sole carbon source.⁴⁸ Interestingly, for *B. fragilis* its ability to cleave fucoses and subsequently incorporate these structures into their own proteins seemed crucial to their fitness in the gut, similar to what was seen in the *H. pylori* experiments.^{49,50} How the relationship between the Bacteroidetes phyla fucosidases and its host differs from the relationship of the host fucosidases with *H. pylori* remains to be elucidated.

In the Actinobacteria phylum, *B. bifidum* ATCC 35914 was isolated as a mucin degrader, and several enzymatic activities involved in mucin degradation were detected in the media when grown in the presence of mucins as sole carbon source.⁵¹ Several more strains of *B. bifidum* have since then been confirmed to be able to degrade mucins.^{52,53} This was followed by the construction of a genomic library of *B. bifidum* in *E. coli* to screen for the ability to hydrolyse specific α -1,2- fucoside linkages which resulted in the isolation of the 1,2- α -L-fucosidase AfC. The cloning and characterization of the AfC enzyme formed a milestone in the field of fucosidase research as it demonstrated the existence of the previously unknown fucosidase family of the GH95 inverting fucosidases.^{54,55}

3. Fucosidases GH29 and GH95

The effect of fucosidases on glycosylation patterns, as well as their temporal and spatial distribution, are only superficially known.⁵⁶ Visualizing where and when specific fucosidases are active will give insight in the way microbiota interact with the epithelial cells and each other.

α -L-Fucosidases catalyse the hydrolysis of α -L-fucoside linkages with the release of L-fucose. Fucosidase-coding genes are widely distributed among the gut microbiota

and are classified within the GH29 (retaining) and GH95 (inverting) families which differ in their enzymatic mechanisms. The two catalytic mechanisms employed by glycosidases to achieve the hydrolysis of a glycosidic cleavage are shown in Figure 5. Both types of enzymes utilise two catalytic carboxylates in their active site.

GH29 retaining glycosidases proceed with overall retention of the anomeric configuration. In 1953, Daniel Koshland proposed that these retaining enzymes act by a double-displacement mechanism and formation of a covalent enzyme intermediate.⁵⁷ The retaining mechanism involves two catalytic carboxylates: one acting as an acid/base catalyst and the other one acting as a nucleophile. In the first step, the catalytic nucleophile of the enzyme performs a nucleophilic attack to a bound glycoside, with the leaving group departure being aided by the catalytic acid/base, leading to the formation of a covalent glycosyl-enzyme intermediate. Subsequently, water enters the active site of the enzyme and gets activated by the catalytic acid/base to act on the trapped glycoside and release the catalytic nucleophile (Figure 5A). The first X-ray study of an α -L-fucosidase was determined using the *Thermotoga maritima* enzyme (TmGH29) in 2004.⁵⁸ Subsequently, a number of crystal structures from other GH29 members were reported, i.e. those from *B. thetaiotaomicron* and *Bifido infantis*.^{59,60} From these crystal structures and related kinetic data several observations could be made. First, the enzymes typically form dimers or higher order oligomers. Second, the catalytic nucleophile in GH29 enzymes seems to be conserved, whereas the catalytic acid/base residue is not.⁶¹ And finally, the substrate specificity of this class of enzymes varies, some enzymes are able to hydrolyse $\alpha(1,2/3/4/6)$ -fucosyl linkages and also work efficiently on the synthetic substrate *p*-nitrophenyl- α -L-fucopyranoside, whereas others seem to act specifically on $\alpha(1,3/4)$ -fucosyl linkages and have poor to no activity on *p*-nitrophenyl- α -L-fucopyranoside. This led to the designation of two subfamilies of GH29 families, where GH29-A have broad substrate specificity and those from the GH29-B sub-family are $\alpha(1,3/4)$ -linkage specific.^{55,62}

Inverting glycosidases are believed to proceed with an overall inversion of the anomeric configuration. The basic carboxylate deprotonates water. The activated hydroxyl then performs an nucleophilic attack on the anomeric carbon, whilst the leaving group is stabilized by the acidic carboxylate. This will result in an inverted configuration of the product compared to that of the substrate glycoside (Figure 5B). However, the catalytic mechanism for GH95 inverting fucosidases is thought to be less standard than typical seen in inverting glycosidases. During the study of *B. bifidum* α -L-fucosidases X-ray crystal structure identified residue Glu⁵⁶⁶ to lie in an appropriate position to act as the catalytic acid as was also further confirmed by mutagenesis experiments, no carboxylate was observed in an appropriate position to act as an catalytic base.⁶³ Rather, Nagae *et al.* concluded that a conserved structural motif creates an electrostatic relay in the enzyme to modulate the pK_a of Asn⁴²³ for its deprotonation of water (Figure 5C).⁶³ Furthermore, this water molecule was shown to be held in the correct position and

orientation for attack on the anomeric carbon by hydrogen bonding interactions by the catalytic bases Asn⁴²³ and Asn^{421, 63}.

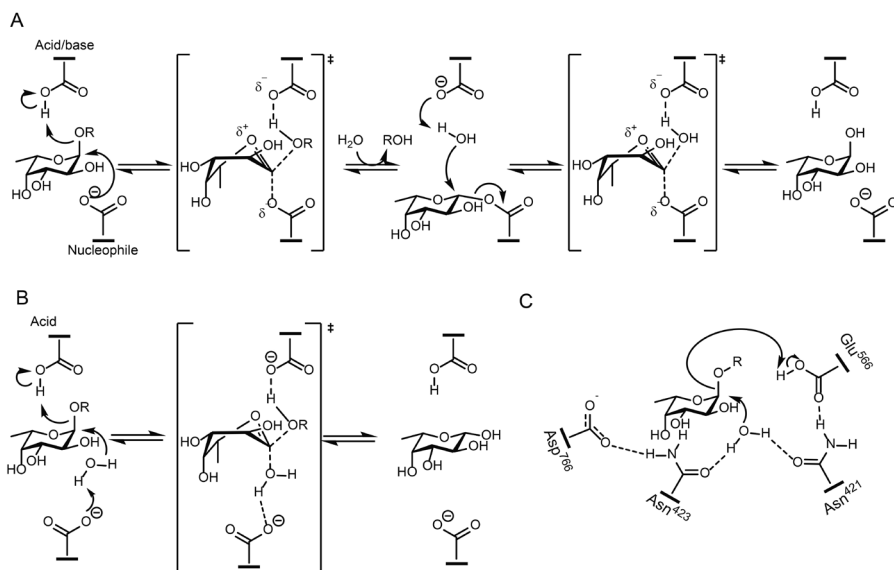


Figure 5 | The retaining (a) and inverting (b) mechanism as employed by α -L-fucosidases. R, a carbohydrate derivative. (c) Proposed catalytic mechanism of the AfcA fucosidase. Hydrogen bonds are depicted by dotted lines.

3.2 Mechanism-based inhibitors and activity-based probes of α -L-fucosidases

Mechanism-based inhibitors have been developed as specific tools to study the interplay of glycosidases on the glycosylation pattern and their downstream effect. The inhibitors follow the mechanism employed by the enzyme as described in the previous paragraph. When a mechanism-based inhibitor is activated in the active site it forms a covalent bond between the inhibitor and the enzyme that is stable over time, thereby inactivating the enzyme. Inactivation is achieved most easily through the formation of a covalent intermediate between the enzyme and the substrate. From this point of view, GH29 retaining fucosidases appear more susceptible to these mechanism-based inhibitors compared to the GH95 inverting fucosidases.

A major group of retaining glycosidase mechanism based inhibitors function by extending the lifetime of the covalent glycosyl-enzyme intermediate.⁶⁴ 2-Deoxy-2-fluoro glycosides were originally introduced by Withers and co-workers.^{65,66} The substitution of a highly electronegative fluorine atom for the C2 hydroxyl group destabilizes the oxocarbenium ion-like states, resulting in a decreased rate for both the formation of the glycosyl-enzyme intermediate and its subsequent hydrolysis (Figure 6A). This was

extended by Vocadlo and Bertozzi who used this design as the basis for activity-based protein profiling (ABPP) to selectively label β -galactosidases in crude cell lysates.⁶⁷ Many other fluorosugar based ABPs have since been synthesized for the labelling of specific glycosidases.^{68–70}

The group of Tatsuna and co-workers was the first to synthesize and annotate cyclophellitol derivatives as mechanism based inhibitors for glycosidases.^{71,72} Cyclophellitol derivatives employ an electrophilic trapping mechanism to target their target. The glycosidase catalytic nucleophile attacks the strained moiety (aziridine, epoxide), leading to irreversible alkylation/acylation of the enzyme (Figure 6B).⁷³ Overkleeft and co-workers extended on this design and provided us with a wide range of ABPs for different retaining glycosidases.^{74–76}

A class of ABPs that can be useful against both retaining as inverting glycosidases relies on the generation of reactive groups. This group is dominated by quinone methide based molecules and has been extensively used to study glycosidases in both biomedical and biotechnical applications.^{77–79} Typically, the enzyme generated quinone methide (QM) features an electrophilic carbon proposed to covalently capture a nucleophile via Michael addition (Figure 6C). Ideally the QM will capture a properly disposed active site residue conserved in the target enzyme resulting in specific protein labelling. However, research has shown that formed QMs might be capable of diffusing out of the active site and label non-specifically.⁸⁰

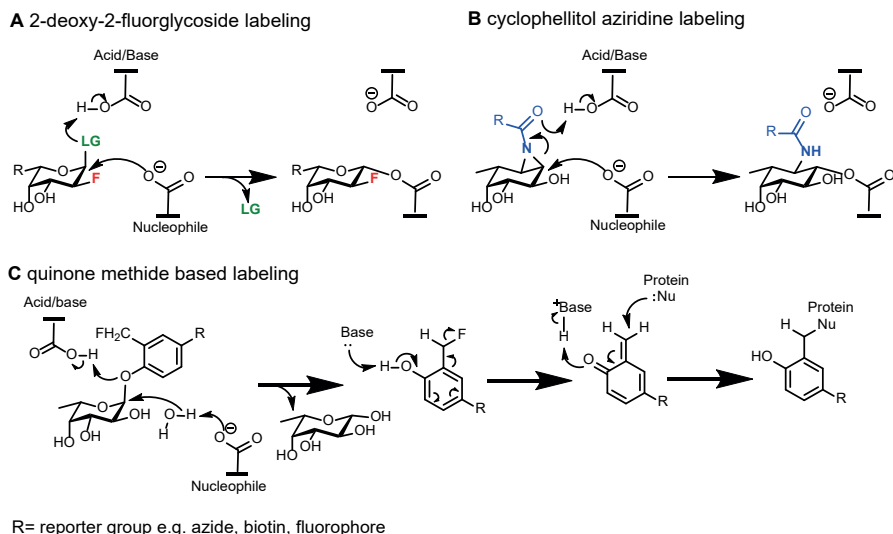


Figure 6 | Mechanism-based inhibition of retaining α -L-fucosidases with (a) 2-deoxy-2-fluoroglycosides (b) cyclophellitol aziridine (c) quinone methide.

3.3 Aim and outline of this thesis

Epithelial cells and surrounding microbiota can interact through the use of specific glycosylation patterns. Specifically, fucosylated glycoproteins have shown to be of biological importance in critical processes to maintain health. The expression of different fucosidases by the numerous bacteria surrounding the epithelial cells can lead to specific alterations of the glycosylation pattern, which can have downstream effects on human intestinal health. Understanding and visualizing these different fucosidases will allow for studying their effect at the human-gut microbiota interface.⁵⁶ In this work, we aimed to expand on the fucosidase ABP library in order to achieve selectivity for different types of fucosidases. In the future this will enable us to study the effect of individual classes of fucosidases at the human-gut microbiota interface.

In **Chapter 2**, we describe our investigation into the influence of fucosidases expressing *Bacteroides* species on specific strains of the pathogenic bacteria *C. jejuni*, known to contain an L-fucose operon. Using PGM or a mucin expressing cell model in combination with the competitive fucosidase inhibitor fuconojirimycin (FNJ), we detected an increased growth and invasion potential for *C. jejuni* in the presence of active fucosidases secreted by *Bacteroides*.

In the human gut, epithelial cells are surrounded by a vast amount of bacteria. These bacteria secrete metabolites, such as butyrate that was previously shown to be capable of inducing the highly O-glycosylated mucins. The Caco-2 cell model is widely used

to study bacterial interactions with mucin *O*-glycans. However, the direct effect of butyrate on *O*-glycan presentation during differentiation of this cell model was not yet investigated. Therefore, in **Chapter 3** we investigated the effect of butyrate on the *O*-glycomics profile of the well-known Caco-2 cell line during its differentiation over a period of 24 days. We employ a PGC nano-LC–ESI-MS/MS method that allows rapid identification and quantification of *O*-glycan alterations. This untargeted in-depth screening of the Caco-2 *O*-glycome proved that there is indeed a specific butyrate dependent *O*-glycan signature when compared to spontaneous differentiation in the absence of butyrate.

After establishing the effect of butyrate on the specific *O*-glycan presentation on Caco-2 cells, we focussed again on the role of L-fucose at the human-gut microbiota interface. As we are interested in studying the identity and activity of gut microbial fucosidases, in **Chapter 4** we expanded on the library of known fucosidases ABPs. We describe the development of a 2-deoxy-2-fluoro fucoside derivative that inactivates and labels α -L-fucosidases by stabilizing the glycosyl-enzyme intermediate. Furthermore, our novel ABP was shown to be active against a GH29 α -L-fucosidases from *B. fragilis* and capable of labeling two other GH29 α -L-fucosidases with different linkage specificity showing its broader utility.

In **Chapter 5**, we report our research on an *O*-fluoromethylphenol fucopyranoside-based probe that is capable of labelling both GH29 and GH95 α -L-fucosidases. Furthermore, we apply bottom-up and top-down MS approaches to identify the specific attachment sites of this probe. Additionally, the possible challenges that may be encountered when using these quinone methide based probes in living systems are discussed.

Finally, **Chapter 6** provides a general discussion about the previous chapters. Furthermore, it discusses several unpublished results and addresses the potential applications of the newly designed ABPs.

References

- 1 F. Sommer and F. Bäckhed, *Nat. Rev. Microbiol.*, 2013, **11**, 227–238.
- 2 L. Wrzosek, S. Miquel, M. L. Noordine, S. Bouet, M. J. Chevalier-Curt, V. Robert, C. Philippe, C. Bridonneau, C. Cherbuy, C. Robbe-Masselot, P. Langella and M. Thomas, *BMC Biol.*, DOI:10.1186/1741-7007-11-61.
- 3 E. Gaudier, L. Forestier, V. Gouyer, G. Huet, R. Julien and C. Hoebler, *Biochem. Biophys. Res. Commun.*, 2004, **325**, 1044–1051.
- 4 D. R. Donohoe, N. Garge, X. Zhang, W. Sun, T. M. O’Connell, M. K. Bunger and S. J. Bultman, *Cell Metab.*, 2011, **13**, 517–526.
- 5 G. M. H. Birchenough, M. E. V. Johansson, J. K. Gustafsson, J. H. Bergström and G. C. Hansson, *Mucosal Immunol.*, 2015, **8**, 712–719.
- 6 A. Ermund, A. Schütte, M. E. V. Johansson, J. K. Gustafsson and G. C. Hansson, *Am. J. Physiol. - Gastrointest. Liver Physiol.*, DOI:10.1152/ajpgi.00046.2013.
- 7 M. E. V. Johansson, *PLoS One*, 2012, **7**, e41009.
- 8 M. E. V. Johansson, J. M. Holmén Larsson and G. C. Hansson, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 4659–4665.
- 9 A. P. Corfield, *Microorganisms*, 2018, **6**, 78.
- 10 L. Arike and G. C. Hansson, *J. Mol. Biol.*, 2016, **428**, 3221–3229.
- 11 S. Neelamegham and L. K. Mahal, *Curr. Opin. Struct. Biol.*, 2016, **40**, 145–152.
- 12 K. W. Moremen, M. Tiemeyer and A. V. Nairn, *Nat. Rev. Mol. Cell Biol.*, 2012, **13**, 448–462.
- 13 H. J. Gabius, *BioSystems*, 2018, **164**, 102–111.
- 14 I. Brockhausen, H. Schachter and P. Stanley, *O-GalNAc Glycans*, 2009.
- 15 C. Robbe, C. Capon, E. Maes, M. Rousset, A. Zweibaum, J. P. Zanetta and J. C. Michalski, *J. Biol. Chem.*, 2003, **278**, 46337–46348.
- 16 J. M. H. Larsson, K. A. Thomsson, A. M. Rodríguez-Piñeiro, H. Karlsson and G. C. Hansson, *Am. J. Physiol. - Gastrointest. Liver Physiol.*, 2013, **305**, G357.
- 17 M. Tong, I. McHardy, P. Ruegger, M. Goudarzi, P. C. Kashyap, T. Haritunians, X. Li, T. G. Graeber, E. Schwager, C. Huttenhower, A. J. Fornace, J. L. Sonnenburg, D. P. B. McGovern, J. Borneman and J. Braun, *ISME J.*, 2014, **8**, 2193–2206.
- 18 M. E. V. Johansson, H. Sjövall and G. C. Hansson, *Nat. Rev. Gastroenterol. Hepatol.*, 2013, **10**, 352–361.
- 19 J. Fu, B. Wei, T. Wen, M. E. V. Johansson, X. Liu, E. Bradford, K. A. Thomsson, S. McGee, L. Mansour, M. Tong, J. M. McDaniel, T. J. Sferra, J. R. Turner, H. Chen, G. C. Hansson, J. Braun and L. Xia, *J. Clin. Invest.*, 2011, **121**, 1657–1666.
- 20 S. J. Chen, X. W. Liu, J. P. Liu, X. Y. Yang and F. G. Lu, *World J. Gastroenterol.*, 2014, **20**, 9468–9475.
- 21 B. O. Schroeder, *Gastroenterol. Rep.*, 2019, **7**, 3–12.
- 22 I. Thiele, J. B. Mason, X. Lin, N. Juge, L. E. Tailford, E. H. Crost and D. Kavanaugh, *Front. Genet.* / www.frontiersin.org, 2015, **6**, 81.
- 23 L. V Hooper, J. Xu, P. G. Falk, T. Midtvedt and J. I. Gordon, *Proc. Natl. Acad. Sci.*, 1999, **96**, 9833–9838.
- 24 S. M. Autieri, J. J. Lins, M. P. Leatham, D. C. Laux, T. Conway and P. S. Cohen, *Infect. Immun.*, 2007, **75**, 5465–5475.

- 25 Y. Goto, S. Uematsu and H. Kiyono, *Nat. Immunol.*, 2016, **17**, 1244–1251.
- 26 P. Wacklin, J. Tuimala, J. Nikkilä, S. Tims, H. Mäkituokko, N. Alakulppi, P. Laine, M. Rajilic-Stojanovic, L. Paulin, W. M. De Vos and J. Mättö, *PLoS One*, DOI:10.1371/journal.pone.0094863.
- 27 C. C. Blackwell, K. Jónsdóttir, M. Hanson, W. T. A. Todd, A. K. R. Chaudhuri, B. Mathew, R. P. Brettle and D. M. Weir, *Lancet*, 1986, **328**, 284–285.
- 28 C. C. Blackwell, K. Jonsdottir, M. F. Hanson and D. M. Weir, *Lancet*, 1986, **328**, 687.
- 29 W. Chaim, B. Foxman and J. D. Sobel, *J. Infect. Dis.*, 1997, **176**, 828–830.
- 30 H. Wu, O. Rebello, E. H. Crost, C. D. Owen, S. Walpole, C. Bennati-Granier, D. Ndeh, S. Monaco, T. Hicks, A. Colvile, P. A. Urbanowicz, M. A. Walsh, J. Angulo, D. I. R. Spencer and N. Juge, *Cell. Mol. Life Sci.*, 2020, **1**, 3.
- 31 G. Vecchio, A. Parascandolo, C. Allocca, C. Ugolini, F. Basolo, M. Moracci, A. Strazzulli, B. Cobucci-Ponzano, M. O. Laukkanen, M. D. Castellone and N. Tsuchida, *Oncotarget*, 2017, **8**, 27075–27092.
- 32 T. C. Cheng, S. H. Tu, L. C. Chen, M. Y. Chen, W. Y. Chen, Y. K. Lin, C. T. Ho, S. Y. Lin, C. H. Wu and Y. S. Ho, *Oncotarget*, 2015, **6**, 21283–21300.
- 33 L. Xu, Z. Li, S. Song, Q. Chen, L. Mo, C. Wang, W. Fan, Y. Yan, X. Tong and H. Yan, *Cancer Sci.*, 2020, **111**, 2284–2296.
- 34 T. W. Liu, C. W. Ho, H. H. Huang, S. M. Chang, S. D. Popat, Y. T. Wang, M. S. Wu, Y. J. Chen and C. H. Lin, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 14581–14586.
- 35 I. Endreffy, G. Björklund, L. Szerafin, S. Chirumbolo, M. A. Urbina and E. Endreffy, *Immunol. Res.*, 2017, **65**, 1025–1030.
- 36 K. Wang, W. Guo, N. Li, J. Shi, C. Zhang, W. Y. Lau, M. Wu and S. Cheng, *Br. J. Cancer*, 2014, **110**, 1811–1819.
- 37 K. M. Ng, J. A. Ferreyra, S. K. Higginbottom, J. B. Lynch, P. C. Kashyap, S. Gopinath, N. Naidu, B. Choudhury, B. C. Weimer, D. M. Monack and J. L. Sonnenburg, *Nature*, 2013, **502**, 96–99.
- 38 M. Stahl, L. M. Friis, H. Nothaft, X. Liu, J. Li, C. M. Szymanski and A. Stintzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 7194–7199.
- 39 J. M. Garber, H. Nothaft, B. Pluvinae, M. Stahl, X. Bian, S. Porfirio, A. Enriquez, J. Butcher, H. Huang, J. Glushka, E. Line, J. A. Gerlt, P. Azadi, A. Stintzi, A. B. Boraston and C. M. Szymanski, *Commun. Biol.*, 2020, **3**, 1–11.
- 40 Y. Luijckx, N. Bleumink, J. Jiang, H. S. Overkleeft, M. Wösten, K. Strijbis and T. Wennekes, *Cell. Microbiol.*, 2020, cmi.13252.
- 41 D. A. Ravcheev and I. Thiele, *Front. Genet.*, 2017, **8**, 111.
- 42 M. Rajilić-Stojanović and W. M. de Vos, *FEMS Microbiol. Rev.*, 2014, **38**, 996–1047.
- 43 A. A. Salyers, J. R. Vercellotti, S. E. H. West and T. D. Wilkins, *Appl. Environ. Microbiol.*, 1977, **33**, 319–322.
- 44 E. C. Martens, H. C. Chiang and J. I. Gordon, *Cell Host Microbe*, 2008, **4**, 447–457.
- 45 E. C. Martens, E. C. Lowe, H. Chiang, N. A. Pudlo, M. Wu, N. P. McNulty, D. W. Abbott, B. Henrissat, H. J. Gilbert, D. N. Bolam and J. I. Gordon, *PLoS Biol.*, 2011, **9**, e1001221.
- 46 J. L. Sonnenburg, J. Xu, D. D. Leip, C. H. Chen, B. P. Westover, J. Weatherford, J. D. Buhler and J. I. Gordon, *Science (80-.)*, 2005, **307**, 1955–1959.
- 47 L. V Hooper, J. Xu, P. G. Falk, T. Midtvedt and J. I. Gordon, *Microbiology*, 1999, **96**, 9833–9838.
- 48 A. M. Roberton and R. A. Stanley, *Appl. Environ. Microbiol.*, 1982, **43**, 325–30.

- 49 M. J. Coyne, B. Reinap, M. M. Lee and L. E. Comstock, *Science* (80-.), 2005, **307**, 1778–1781.
- 50 C. M. Fletcher, M. J. Coyne, O. F. Villa, M. Chatzidaki-livanis and L. E. Comstock, *Cell*, 2009, **137**, 321–331.
- 51 L. C. Hoskins, M. Agustines, W. B. McKee, E. T. Boulding, M. Kriaris and G. Niedermeyer, *J. Clin. Invest.*, 1985, **75**, 944–953.
- 52 P. Ruas-Madiedo, M. Gueimonde, M. Fernández-García, C. G. De Los Reyes-Gavilán and A. Margolles, *Appl. Environ. Microbiol.*, 2008, **74**, 1936–1940.
- 53 F. Turrone, F. Bottacini, E. Foroni, I. Mulder, J. H. Kim, A. Zomer, B. Sánchez, A. Bidossi, A. Ferrarini, V. Giubellini, M. Delledonne, B. Henrissat, P. Coutinho, M. Oggioni, G. F. Fitzgerald, D. Mills, A. Margolles, D. Kelly, D. Van Sinderen and M. Ventura, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 19514–19519.
- 54 T. Katayama, A. Sakuma, T. Kimura, Y. Makimura, J. Hiratake, K. Sakata, T. Yamanai, H. Kumagai and K. Yamamoto, *J. Bacteriol.*, 2004, **186**, 4885–4893.
- 55 H. Ashida, A. Miyake, M. Kiyohara, J. Wada, E. Yoshida, H. Kumagai, T. Katayama and K. Yamamoto, *Glycobiology*, 2009, **19**, 1010–1017.
- 56 F. C. Pereira and D. Berry, *Environ. Microbiol.*, 2017, **19**, 1366–1378.
- 57 D. E. Koshland, *Biol. Rev.*, 1953, **28**, 416–436.
- 58 G. Sulzenbacher, C. Bignon, T. Nishimura, C. A. Tarling, S. G. Withers, B. Henrissat and Y. Bourne, *J. Biol. Chem.*, 2004, **279**, 13119–13128.
- 59 L. Guillotin, P. Lafite and R. Daniellou, *Biochemistry*, 2014, **53**, 1447–1455.
- 60 D. A. Sela, D. Garrido, L. Lerno, S. Wu, K. Tan, H. J. Eom, A. Joachimiak, C. B. Lebrilla and D. A. Mills, *Appl. Environ. Microbiol.*, 2012, **78**, 795–803.
- 61 F. A. Shaikh, A. Lammerts Van Bueren, G. J. Davies and S. G. Withers, *Biochemistry*, 2013, **52**, 5857–5864.
- 62 H. Sakurama, E. Tsutsumi, H. Ashida, T. Katayama, K. Yamamoto and H. Kumagai, *Biosci. Biotechnol. Biochem.*, 2012, **76**, 1022–1024.
- 63 M. Nagae, A. Tsuchiya, T. Katayama, K. Yamamoto, S. Wakatsuki and R. Kato, *J. Biol. Chem.*, 2007, **282**, 18497–18509.
- 64 L. Wu, Z. Armstrong, S. P. Schröder, C. de Boer, M. Artola, J. M. Aerts, H. S. Overkleeft and G. J. Davies, *Curr. Opin. Chem. Biol.*, 2019, **53**, 25–36.
- 65 S. G. Withers, K. Rupitz and I. P. Street, *J. Biol. Chem.*, 1988, **263**, 7929–7932.
- 66 S. J. Williams and S. G. Withers, *Carbohydr. Res.*, 2000, **327**, 27–46.
- 67 D. J. Vocadlo and C. R. Bertozzi, *Angew. Chemie - Int. Ed.*, 2004, **43**, 5338–5342.
- 68 C. P. Phenix, B. P. Rempel, K. Colobong, D. J. Doudet, M. J. Adam, L. A. Clarke and S. G. Withers, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 10842–10847.
- 69 O. Hekmat, Y. W. Kim, S. J. Williams, S. He and S. G. Withers, *J. Biol. Chem.*, 2005, **280**, 35126–35135.
- 70 K. A. Stubbs, A. Scaffidi, A. W. Debowski, B. L. Mark, R. V. Stick and D. J. Vocadlo, *J. Am. Chem. Soc.*, 2008, **130**, 327–335.
- 71 K. Tatsuta, Y. Niwata, K. Umezawa, K. Toshima and M. Nakata, *J. Antibiot. (Tokyo)*, 1991, **44**, 456–458.
- 72 K. Tatsuta, Y. Niwata, K. Umezawa, K. Toshima and M. Nakata, *Tetrahedron Lett.*, 1990, **31**, 1171–1172.

- 73 K. Y. Li, J. Jiang, M. D. Witte, W. W. Kallemeyjn, H. Van Den Elst, C. S. Wong, S. D. Chander, S. Hoogendoorn, T. J. M. Beenakker, J. D. C. Codée, J. M. F. G. Aerts, G. A. Van Der Marel and H. S. Overkleeft, *European J. Org. Chem.*, 2014, **2014**, 6030–6043.
- 74 J. Jiang, W. W. Kallemeyjn, D. W. Wright, A. M. C. H. Van Den Nieuwendijk, V. C. Rohde, E. C. Folch, H. Van Den Elst, B. I. Florea, S. Scheij, W. E. Donker-Koopman, M. Verhoek, N. Li, M. Schürmann, D. Mink, R. G. Boot, J. D. C. Codée, G. A. Van Der Marel, G. J. Davies, J. M. F. G. Aerts and H. S. Overkleeft, *Chem. Sci.*, 2015, **6**, 2782–2789.
- 75 L. I. Willems, J. Jiang, K. Y. Li, M. D. Witte, W. W. Kallemeyjn, T. J. N. Beenakker, S. P. Schröder, J. M. F. G. Aerts, G. A. Van Der Marel, J. D. C. Codée and H. S. Overkleeft, *Chem. - A Eur. J.*, 2014, **20**, 10864–10872.
- 76 L. Wu, J. Jiang, Y. Jin, W. W. Kallemeyjn, C. L. Kuo, M. Artola, W. Dai, C. Van Elk, M. Van Eijk, G. A. Van Der Marel, J. D. C. Codée, B. I. Florea, J. M. F. G. Aerts, H. S. Overkleeft and G. J. Davies, *Nat. Chem. Biol.*, 2017, **13**, 867–873.
- 77 M. Nandakumar, Y. L. Hsu, J. C. Y. Lin, C. Lo, L. C. Lo and C. H. Lin, *ChemBioChem*, 2015, **16**, 1555–1559.
- 78 Y. L. Hsu, M. Nandakumar, H. Y. Lai, T. C. Chou, C. Y. Chu, C. H. Lin and L. C. Lo, *J. Org. Chem.*, 2015, **80**, 8458–8463.
- 79 L. M. Chauvigné-Hines, L. N. Anderson, H. M. Weaver, J. N. Brown, P. K. Koech, C. D. Nicora, B. A. Hofstad, R. D. Smith, M. J. Wilkins, S. J. Callister and A. T. Wright, *J. Am. Chem. Soc.*, 2012, **134**, 20521–20532.
- 80 J. Lenger, M. Schröder, E. C. Ennemann, B. Müller, C. H. Wong, T. Noll, T. Dierks, S. R. Hanson and N. Sewald, *Bioorganic Med. Chem.*, 2012, **20**, 622–627.

Chapter 2

***Bacteroides fragilis* fucosidases facilitate growth and invasion of *Campylobacter jejuni* in the presence of mucins**

Yvette M.C.A. Luijkx, Nancy M.C. Bleumink, Jianbing Jiang, Herman S. Overkleeft, Marc M.S.M. Wösten, Karin Strijbis, Tom Wennekes

Adapted from: Cell. Microbiol., 2020, cmi.13252

Abstract

The enteropathogenic bacterium *Campylobacter jejuni* was considered to be non-saccharolytic, but recently it emerged that L-fucose plays a central role in *C. jejuni* virulence. Half of *C. jejuni* clinical isolates possess an operon for L-fucose utilization. In the intestinal tract, L-fucose is abundantly available in mucin O-linked glycan structures, but *C. jejuni* lacks a fucosidase enzyme essential to release the L-fucose. We set out to determine how *C. jejuni* can gain access to these intestinal L-fucosides. Growth of the *fuc*⁺ *C. jejuni* strains 129108 and NCTC 11168 increased in the presence of L-fucose while fucose permease knockout strains did not benefit from additional L-fucose. With fucosidase assays and an activity-based probe we confirmed that *Bacteriodes fragilis*, an abundant member of the intestinal microbiota, secretes active fucosidases. In the presence of mucins, *C. jejuni* was dependent on *B. fragilis* fucosidase activity for increased growth. *C. jejuni* invaded Caco-2 intestinal cells that express complex O-linked glycan structures that contain L-fucose. In infection experiments, *C. jejuni* was more invasive in the presence of *B. fragilis* and this increase is due to fucosidase activity. We conclude that *C. jejuni fuc*⁺ strains are dependent on exogenous fucosidases for increased growth and invasion.

Introduction

Enteropathogenic bacteria that infect the human intestinal tract have the ability to proliferate and invade within the intestinal niche that is dominated by its mucus layer and associated O-glycans.^{1,2} The intestinal mucus layer consists of two distinct layers that together form a physical barrier protecting the underlying epithelium against invasive pathogens and the residing intestinal microbiota.^{3–5} Under healthy conditions the microbiota are associated with the outer mucus layer while the inner layer is relatively impenetrable by bacteria.⁶ At the host-microbe interface, the highly glycosylated mucins provide the resident microbiota with nutrients that are primarily derived from its O-glycans.⁷ Invading enteropathogenic bacteria have to compete for nutrients and space at this interface in order to proliferate and cause infection.

Mucin O-linked glycans contribute 80% of the total dry weight of intestinal mucins and are made up of the monosaccharides fucose, galactose, N-acetylgalactosamine, N-acetylneuraminic acid (sialic acid), and N-acetylglucosamine.^{8,9} Numerous bacterial glycosidases have been identified that liberate monosaccharides from the complex O-glycan structure (Thiele *et al.*, 2015, PMID:24270786). It has been predicted by bioinformatics that up to 40% of all commensal bacterial encode for carbohydrate-degrading enzymes (glycosidases) and monosaccharide-utilizing enzymes.¹⁰

One bacterial phylum that is well-equipped to degrade polysaccharides are the Bacteroidetes that make up almost half of the bacteria found in the intestine.¹¹ Genomic and proteomic analysis of *Bacteroides thetaiotaomicron* demonstrated that this bacterium possesses over 280 glycosidases, of which 11% are located on the outer membrane or released extracellularly.^{11,12} Sialidases and fucosidases are the glycosidases that target the two major terminal epitopes found on mucin O-linked glycans, sialic acid and fucose. In the human intestine the density of the fucosylated O-glycans decreases from ileum to colon while sialylated O-glycans show the reversed pattern.⁸ Both fucose and sialic acid are reported in literature to be correlated with the ability of pathogenic bacteria to thrive within the gut.^{13–15} Enteropathogenic bacteria need to compete for nutrients with resident commensals to gain a foothold in our gut and it is thus important to understand how they access mucin-derived monosaccharides to establish infection.

The enteropathogen *Campylobacter jejuni* is a microaerophilic Gram-negative, spiral shaped bacterium with bipolar flagella. This species is the leading cause of bacterial gastroenteritis in the developing world.^{16–18} *C. jejuni* isolates were for a long time considered to be non-saccharolytic, meaning that these bacteria do not metabolize sugars.^{19,20} However, both Muraoka and Zhang and Stahl *et al.*, have reported that certain strains of *C. jejuni* have the ability to utilize L-fucose.^{21,22} They showed that two commonly studied *C. jejuni* strains, NCTC 11168 and RM1221 have a growth benefit in the presence of L-fucose. Previous studies have identified potential similarities between

the L-fucose breakdown pathway of *C. jejuni* and the plant pathogen *Xanthomonas campestris*.²³ A recent study elucidated the L-fucose breakdown pathway in the *C. jejuni* NCTC 11168 *fuc+* strain by solving the structure of a putative dehydrogenase, FucX, that can reduce L-fucose and D-arabinose *in vitro*.²⁴

Besides the effects of L-fucose breakdown for basic metabolism, L-fucose utilization has a broader impact on *C. jejuni* biology. Transcriptomics analysis of *C. jejuni fuc+* strains showed a large fold change in transcript abundancies upon addition of L-fucose with 74 transcripts up-regulated and another 52 down-regulated.²² For example, up-regulation was seen for the immunoreactive *cstA* (*cj0917c*), a carbon starvation protein A homolog.^{22,25} However, explanations for these up- and down- regulations are not immediately clear. Furthermore, a recent metabolomics study has demonstrated that *C. jejuni fuc+* strains have an adaptive metabolome that changes in the presence of L-fucose.²³ Metabolites dependent on L-fucose, such as thiazolidine-containing metabolites, could be detected that demonstrate the activation of metabolic pathways generating bio active compounds in *C. jejuni*.²³

The fucose operon is not conserved universally among *C. jejuni* strains, but its presence has been linked to hyper invasiveness in *in vitro* virulence and transposon mutagenesis.^{26,27} The capacity of *C. jejuni fuc+* strains to utilize L-fucose correlates with colonization and pathogenicity advantages in neonatal piglet model.²² Interestingly, the *C. jejuni* genomes that have been sequenced so far lack fucosidases that would be necessary to release L-fucose from host mucins. A recent publication verified a lack of *C. jejuni* fucosidases and showed increased growth of *C. jejuni fuc+* strains in the presence of fucosidases secreted by commensal bacteria.²⁴ Complementary to this finding we want to determine the effect of fucosidase activity of residing commensal bacteria on *C. jejuni fuc+* strains hyper invasiveness in the presence of mucins.

Fucosidases can be classified in two different glycoside hydrolase families. GH29 fucosidases have a retaining mechanism by which they form a temporary covalent bond with fucosyl residue, while members of the GH95 family employ an inverting mechanism. GH29 enzymes (EC 3.2.1.111 & EC 3.2.1.51) have a broad specificity to α 1,2/3/4/6-fucosidic linkages, whereas GH95 enzymes (EC 3.2.1.63) are specific to α 1,2-fucosidic linkages. A previous proteomics study on *B. fragilis* showed the presence of two putative secreted GH29 α -L-fucosidases.²⁸ In our study we utilize an activity-based probe (ABP) that mimics a α -L-fucosyl residue and can detect catalytically active GH29 fucosidases.²⁹

To investigate the dependence of *C. jejuni* on exogenous fucosidases activity and its implications for growth and virulence, we took an interdisciplinary approach by combining microbiology with an activity-based probe (ABP), competitive fucosidase inhibitors and infection assays. With the use of the ABP we were capable of visualizing

active GH29 fucosidases secreted by *B. fragilis*. Using this diverse approach, we demonstrated that fucosidases secreted by commensal *Bacteroides fragilis* increased growth and invasion of the *C. jejuni fuc+* strain in the presence of mucin O-linked glycans. Our data shed light on nutrient competition in the intestinal tract and the contribution of commensal glycosidases to the virulence of enteropathogens.

Results

The hyperinvasive *C. jejuni* 108 strain contains the genomic island required for fucose utilization

The *fuc+* operon containing genes *cj0480c* to *cj0490*, has been identified in *C. jejuni* NCTC 11168 and RM1221 to be required for L-fucose utilization.²² We sequenced the hyperinvasive *C. jejuni* 129108 (108) strain and identified 11 genes to be homologous to the *cj0480-cj0490* gene cluster of NCTC 11168 with a sequence similarity of 98.91%. Figure 1 shows the schematic representation of the *fuc+* operon. The genes encoded by the *C. jejuni fuc+* operon are predicted to include a transcriptional regulator (FucR), a synthase (*dapA*), a dehydratase (*uxaA'*), two major facilitator superfamily transporters (Cj0484 and FucP), two dehydrogenases (FucX and Cj0489), a hydrolase (Cj0487), and a mutarotase (Cj0488).^{22,24} Cj0486 is homologous to fucose permeases found in other bacteria and was previously shown to be an essential component of the active L-fucose assimilation pathway in *C. jejuni* NCTC 11168.²² The predicted Cj0486 gene product in *C. jejuni* 108 is 99% identical to its NCTC 11168 homologue. Based on its sequence, we predict that the hyperinvasive *C. jejuni* 108 is a *fuc+* strain that has the ability to scavenge and metabolize L-fucose.

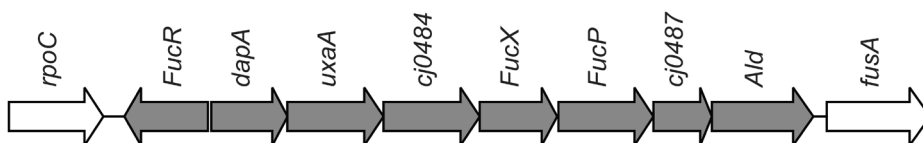


Figure 1 | Schematic representation of the *fuc+* operon as is appears in *C. jejuni* NCTC 11168 and 1219108.

L-fucose increases growth of *C. jejuni* 108

We next investigated the effect of L-fucose on growth of the *C. jejuni* 108 strain. Strains 108 and NCTC 11168 were grown in DMEM with and without L-fucose. In the presence of L-fucose, both strains reached a higher final optical density, but the growth increase was most pronounced in the 108 strain (Fig. 2A). In contrast with L-fucose, no significant increase in growth of *C. jejuni* 108 was seen in the presence of sialic acid (Fig. 2B). We

generated deletion strains for the fucose permease *cj0486* for both the 11168 and 108 strains and tested their ability to grow on L-fucose. Growth of the mutant strains was similar with or without addition of L-fucose, indicating that both mutants lost their ability to utilize L-fucose (Fig. 2C). These results demonstrate that the hyperinvasive *C. jejuni* 108 strain contains a pathway for the uptake and metabolism of L-fucose and that L-fucose confers a growth benefit.

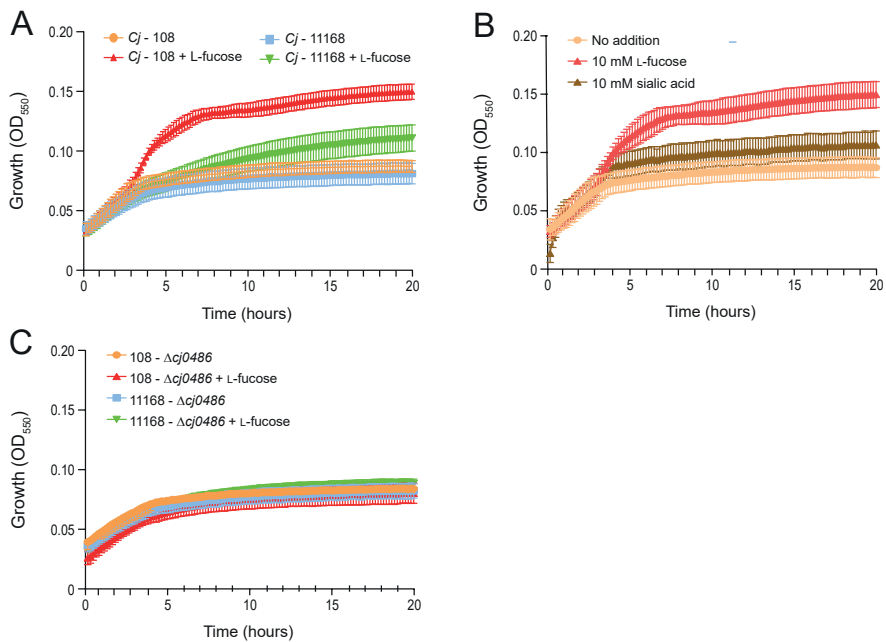


Figure 2 | Growth of *C. jejuni* is enhanced in the presence of L-fucose. Growth curves in DMEM of (a) *C. jejuni* strains, NCTC 11168 and 108, supplemented with either L-fucose or sialic acid, (b) *C. jejuni* supplemented with 10 mM sialic acid or 10 mM L-fucose, (c) *C. jejuni* NCTC 11168 and 108 Δ *cj0486* mutants with and without 10 mM L-fucose. All experiments were performed three times, error bars show SEM.

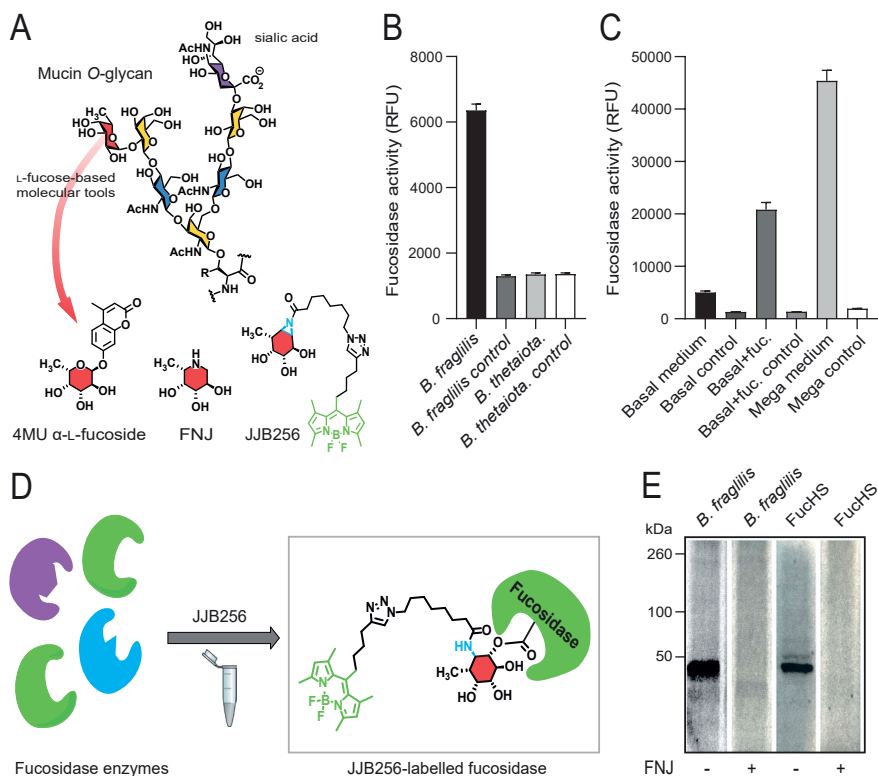


Figure 3 | *Bacteroides fragilis* secretes active GH29 fucosidases. (a) Mucin O-glycan structure and overview of the L-fucose-based molecular tools used in this study. 4MU-fuc as fluorogenic substrate, reversible inhibitor FNJ, activity-based probe JJB256 for in-gel visualisation and quantification. Fucose (red), sialic acid (purple), N-acetylgalactosamine (yellow), N-acetylglucosamine (blue). Fucosidase activity as determined by 4-methylumbelliferyl- α -L-fucopyranoside (4MU α -L-fucopyranoside) assay in (b) supernatant fraction of *B. fragilis* and *B. thetaiotaomicron* grown on basal medium, negative controls show the background level of the heat-inactivated samples, (c) *Bacteroides fragilis* supernatant after growth in basal medium, basal medium with 10 mM L-fucose or the rich Mega medium. Fucosidase activity was expressed in relative fluorescence units (RFU) and negative controls show the background level of the heat-inactivated samples. All experiments were performed three times, error bars show SEM. (d) Schematic representation of interaction of the JJB256 activity-based probe (ABP) with a target GH29 fucosidase from a complex mixture of enzymes. (e) Detection of fucosidase activity with ABP JJB256 on FucHS and supernatant of *B. fragilis* by in-gel fluorescence. As a negative control L-fuconojirimycin (FNJ) was taken along.

Detection of fucosidase activity of commensal *Bacteroides fragilis* using chemical tools

Our results demonstrate that *C. jejuni fuc+* strains can utilize the monosaccharide L-fucose, but the release of L-fucose from complex mucin O-linked glycans requires extracellular fucosidase activity (mucin O-glycan Fig. 3A). We hypothesize that *C. jejuni*

fuc+ strains are dependent on fucosidase activity of residing commensals such as *Bacteroides* species. We selected *Bacteroides fragilis* and *Bacteroides thetaiotaomicron* for our experiments to induce extracellular fucosidases and detect their activity with a set of molecular tools (the structures of the tools used in our assays are depicted in Fig. 3A). *B. fragilis* and *B. thetaiotaomicron* were grown anaerobically in basal medium (TYG) and activity of secreted fucosidases was measured by the breakdown of the fluorogenic substrate 4-methylumbelliferyl- α -L-fucopyranoside. Fucosidase activity was detectable in secreted fractions of *B. fragilis* grown under these conditions, but not in the supernatants of *B. thetaiotaomicron* or heat-inactivated supernatant of *B. fragilis* (Fig. 3B). Next, we determined if addition of L-fucose or growth in Mega medium, which is specifically optimized for growth of intestinal anaerobic bacteria, would increase secreted fucosidase activity of *B. fragilis*. Addition of L-fucose and growth in Mega medium enhanced secreted fucosidase activity (Fig. 3C), which is in line with previous observations.²⁸

In a previous study, proteomics analysis showed the presence of secreted GH29 fucosidases in this specific *B. fragilis* strain.²⁸ To confirm the nature of the fucosidase enzymes secreted by *B. fragilis* in our experimental set up, we used an activity-based fucosidase probe (JJB256) that was previously synthesized and applied by the Overkleeft group.²⁹ When JJB256 is bound by a catalytically active GH29 fucosidase its reactive warhead (aziridine, blue, Fig. 3D) will react to form a stable covalent bond within the fucosidase active site. The fluorophore that is attached to JJB256 (green, Fig. 3D) allows for visualization of the labeled fucosidases by in-gel fluorescence (Fig. 3E). We performed labeling experiments with JJB256 and found that it efficiently labeled the control *Homo sapiens* GH29 fucosidase (FuchS). In addition, we could detect positive labeling of fucosidases in the supernatant fraction of *B. fragilis* (Fig. 3E). The prominent band around 50 kDa is in agreement with the predicted molecular weight of *B. fragilis* secreted fucosidases in the CAZy database (<http://www.cazy.org/GH29.html>). Pre-incubation with 100 μ M of the competitive fucosidase inhibitor L-fuconojirimycin (FNJ) blocked labelling of the major 50 kDa bands supporting the presence of active GH29 fucosidases in the FuchS and *B. fragilis* supernatant fractions. We conclude that *Bacteroides fragilis* secretes active fucosidases of the GH29 family during anaerobic growth in Mega medium.

***C. jejuni* is dependent on *B. fragilis* fucosidases for growth on mucin**

We next investigated if *C. jejuni* 108 can benefit from exogenous fucosidase activity for growth on mucin O-linked glycans. Porcine gastric mucin (PGM) was pretreated with or without purified fucosidase FuchS and subsequently *C. jejuni* 108 was added and incubated for 24 hours. Pretreatment of the mucin with the fucosidase enzyme resulted in a significant increase in *C. jejuni* colony forming units (CFUs) compared to non-treated PGM (Fig 4A). Pretreatment of PGM with sialidase did not confer a significant growth

benefit for *C. jejuni* 108 (Fig 4B), which is in line with our earlier observations in the growth assays with sialic acid.

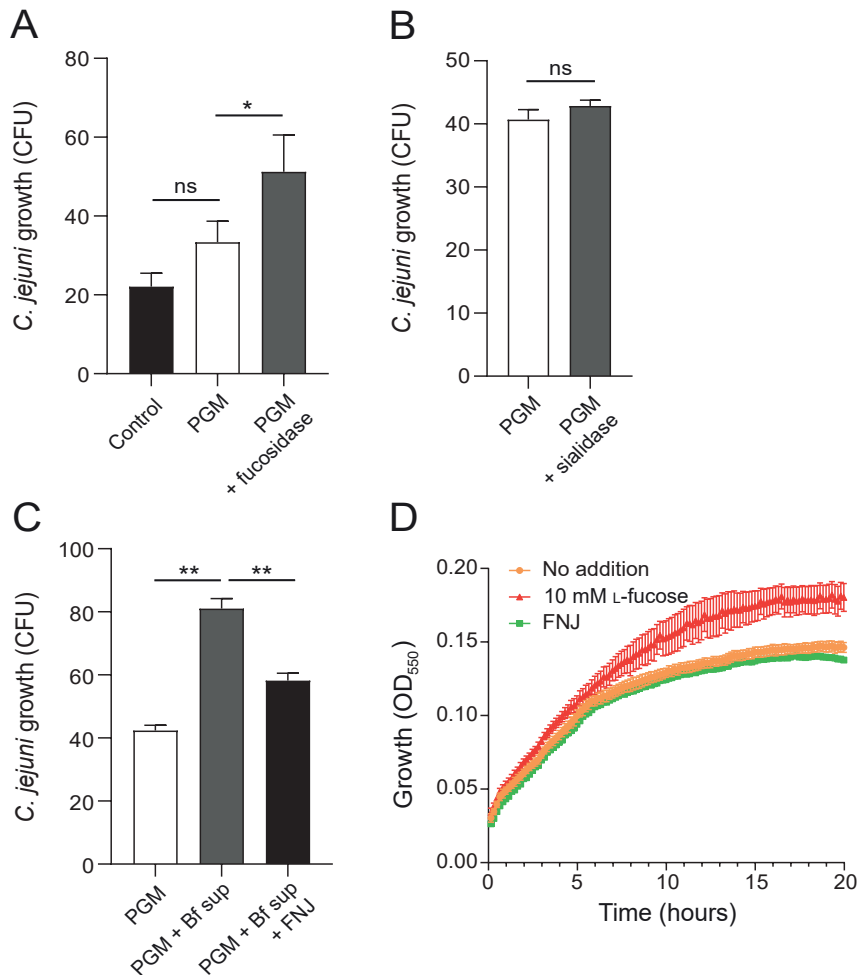


Figure 4 | *Campylobacter jejuni* 108 requires exogenous fucosidases for growth on mucin. Growth of *C. jejuni* 108 on pig gastric mucin (PGM) without or with (a) recombinant fucosidase FUCHS, (b) recombinant sialidase, (c) *B. fragilis* supernatant in the presence of 100 μ M competitive fucosidase inhibitor FNJ. Data shown in bar graphs represent means \pm SEM ($n = 3$ independent experiments); * $p < .05$; ** $p < .01$. (d) Growth curve of *C. jejuni* 108 in DMEM medium supplemented with 10 mM L-fucose or 100 μ M FNJ. Experiments were performed three times, error bars show SEM.

Next, we set out to investigate the effects of the more complex *B. fragilis* supernatant. We normalized fucosidase activity in the supernatant fraction to the previously used FuchS activity by comparing their in-gel fluorescent signals with probe JJB256 using

ImageJ software. PGM was pre-treated with the concentrated *B. fragilis* supernatant fraction and added to the *C. jejuni* 108 culture. *C. jejuni* growth was significantly increased in the presence of the pre-treated PGM compared to the untreated PGM (Fig. 4C). Addition of 100 μ M FNJ to the *B. fragilis* supernatant fraction resulted in a significant decrease in *C. jejuni* growth on PGM, demonstrating that the increase in *C. jejuni* growth is due to *B. fragilis* fucosidase activity (Fig. 4C). As a control we also investigated the effect of FNJ on the growth of *C. jejuni* 108. Growth curves of *C. jejuni* 108 were similar in the presence and absence of FNJ, indicating that this inhibitor does not directly impact growth (Figure 4D). Together, these results demonstrate that for growth on glycosylated mucin, *C. jejuni* 108 is dependent on secreted fucosidases from other species.

Effects of fucosidase activity on *C. jejuni* 108 invasion into intestinal epithelial cells

Fucosylated mucins are also expressed on the apical surface of intestinal epithelial cells and could be possible targets for *C. jejuni* adhesion and invasion as previously suggested.^{30,31} We investigated the presence of fucosylated mucins on the surface of the intestinal epithelial Caco-2 cells after 5 days of differentiation. Immunofluorescence with UEA-I lectin (H type 2) and the α H(O)I antibody (H type 1) showed the presence of H type 2 fucosylated structures evenly distributed over the cell surface and membrane-bound mucin associated H-type 1 fucosylation as patches (Fig. 5A).

Microaerophilic growth conditions are optimal for *C. jejuni* and the importance of mimicking *in vivo* conditions for *C. jejuni* has previously been demonstrated.³² Therefore, we performed infection assays under microaerophilic conditions that are favorable to *C. jejuni* and more representative of the intestinal epithelial interface. A *C. jejuni* 108 strain expressing mCherry was used to allow imaging of invasion by confocal microscopy. Confluent Caco-2 cells were differentiated for 5 days and infected with *C. jejuni* 108 for 3.5 h under microaerophilic conditions. We observed that *C. jejuni* 108 predominantly invaded L-fucose positive cells (Fig. 5B). Multiple Z-stacks of the image were collected, and an orthogonal view showed apical fucosylated mucins with *C. jejuni* 108 located underneath (Fig. 5C).

To quantify invasion of *C. jejuni* into intestinal cells under different conditions, we performed gentamicin protection assays with Caco-2 cells. Addition of 10 mM L-fucose significantly increased the number of intracellular *C. jejuni* as measured by CFUs after lysis of the Caco-2 cells (Fig. 6A). This increased invasion might be in part attributed to the growth-stimulating effect of L-fucose on *C. jejuni* 108 that we observed previously. Next, we added FucHS during the infection assay to release L-fucose from the mucin O-glycans on the Caco-2 cell surface. Addition of exogenous fucosidases resulted in a significant increase in intercellular *C. jejuni* 108 (Fig. 6B). Based on these results we conclude that free extracellular fucose increases *C. jejuni* invasion into intestinal epithelial cells.

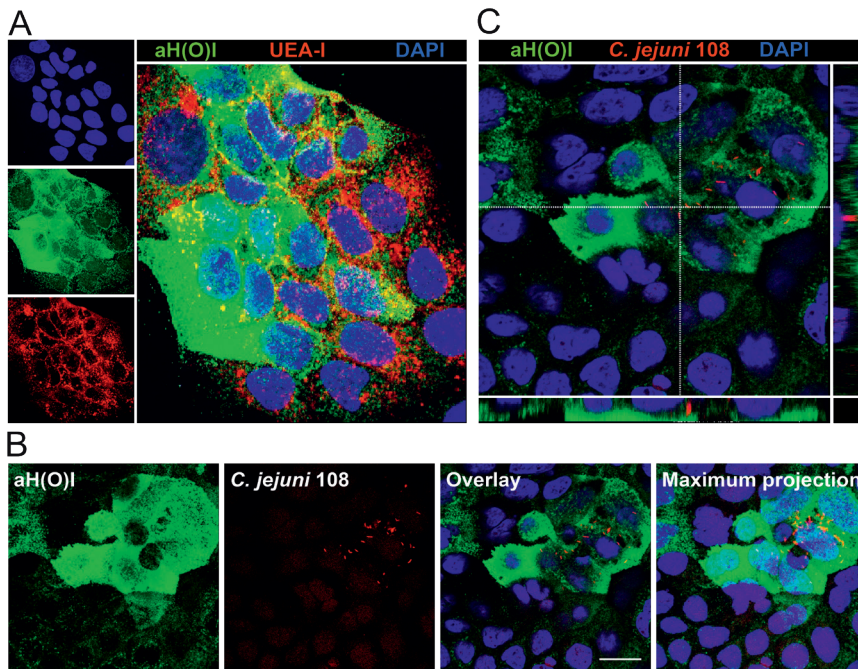


Figure 5 | Presence of L-fucosides on the apical surface of Caco-2 epithelial cells correlates with *Campylobacter jejuni* 108 invasion. (a) Immunofluorescence confocal microscopy image of confluent Caco-2 cells stained for L-fucosides with UEA-I lectin (α 1,2 fucose H-type2, red), or aH(O)I antibody (α 1,2 fucose H-type 1, green) and nuclei (DAPI, blue). (b, c) Immunofluorescence confocal microscopy image of confluent Caco-2 cells infected with *C. jejuni* 108 (mCherry, red) stained with aH(O)I antibody for α 1,2 fucose H-type 1 (green) and nuclei (DAPI, blue) demonstrating invasion of *C. jejuni* into Caco-2 cells with apical L-fucose. White scale bars represent 20 μ m.

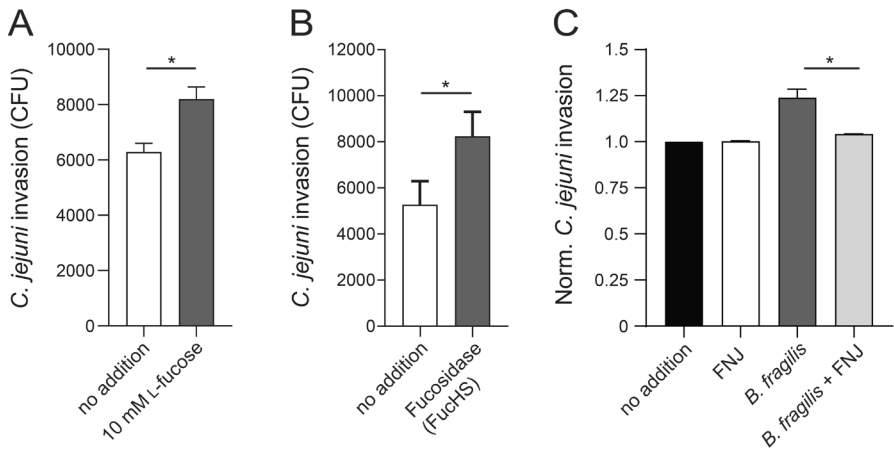


Figure 6 | Increased invasion of *Campylobacter jejuni*108 into intestinal epithelial cells in the presence of L-fucose. *Campylobacter jejuni*108 gentamicin survival assays in Caco-2 epithelial cells performed under microaerophilic conditions. Quantification of *C. jejuni* 108 invasion into Caco-2 cells without or with 10 mM L-fucose (a) and with or without fucosidase FUCHS (b). *Bacteroides fragilis* in the presence or absence of the fucosidase inhibitor FNJ (c). Data shown in bar graphs represent means \pm SEM (n = 3 independent experiments); * p < .05.

As established above, *C. jejuni* benefits from fucosidase secreted by *B. fragilis* in the presence of fucosylated mucins. However, during infection *in vivo*, *Bacteroides* species also compete for the liberated L-fucose.³³ Therefore, our final experiment was to determine the effect of live *B. fragilis* bacteria on *C. jejuni* 108 invasion. *B. fragilis*, grown anaerobically in Mega medium, was added during the infection assay with Caco-2 cells at a MOI 1:10 (*B. fragilis* : *C. jejuni* 108). To determine the impact of fucosidases on *C. jejuni* invasion, we conducted the experiment in the presence or absence of the competitive fucosidase inhibitor FNJ. The presence of *B. fragilis* boosted *C. jejuni* 108 invasion and this effect was abolished in the presence of the competitive inhibitor FNJ (Fig. 6C). We conclude that the increased invasion of *C. jejuni* 108 was caused by fucosidase activity present during the co-culture with *B. fragilis*. Furthermore, a previous study that reported on proteomics analysis of this *B. fragilis* strain showed only secreted GH29 fucosidases activity and no cell membrane bound fucosidases.²⁸ Taken together, these results support our hypothesis that locally liberated L-fucose by secreted fucosidases from other species can increase growth and invasion of *fuc+* *C. jejuni* strains at the intestinal epithelial interface.

Discussion

When *C. jejuni* invades the mucosal lining of the intestinal epithelium it has to compete with residing intestinal microbiota for nutrients and space.³⁴ Although *C. jejuni* prefers growth on amino acids, some hyperinvasive strains possess the ability to metabolize L-fucose, which is an abundant terminal component of mucin O-glycans that cover the intestinal epithelium.^{26,27} The ability of *C. jejuni fuc+* strains to metabolize L-fucose confers a competitive advantage in infection models.²² However, as *C. jejuni* lacks endogenous fucosidases it does not have the capacity to release L-fucose from mucin O-glycans. Our results here show that secreted fucosidases of a commensal, *Bacteroides fragilis*, facilitate enhanced growth of the hyperinvasive *C. jejuni fuc+* strain 108 on glycosylated mucins (Figure 7). Furthermore, we used activity-based protein profiling (ABPP) and chemical competitive inhibitors to demonstrate the crucial contribution of these exogenous fucosidases to the increased invasion by *C. jejuni* 108 into intestinal epithelial cells. Our findings complement two recent publications on the topic of *C. jejuni fuc+* strains that investigated nutrient scavenging by *C. jejuni fuc+* strains and demonstrated how the presence of glycoproteins in human milk affects the selection of these specific strains.^{24,35} Furthermore, it is interesting to speculate that our data are in line with a previous finding that individuals with higher proportions of *Bacteroides* species are more susceptible to *C. jejuni* infections.³⁶

Bacteroides species are known to tightly regulate the secretion of their extracellular fucosidases, which levels often appear lower *in vitro* compared to *in vivo* experiments.³⁷ ABPP is an established and powerful technique to label and detect catalytically active enzymes in their native environment.³⁸ Our successful application of chemical probe JJB256 to label previously putative retaining GH29 α -L-fucosidases in the secretome of *B. fragilis* highlights the possible further application of this probe and potential future derivatives.²⁹ One such application could be screens for secreted GH29 fucosidases in both microbiota and pathogens.

The L-fucose cross feeding that we observe between the commensal *Bacteroides fragilis*, and invading pathogen *Campylobacter jejuni* is a strategy encountered more often among enteropathogenic bacteria. The virulence genes of Enterohemorrhagic *E. coli* (EHEC) are regulated by the fucose-activated FusKR signaling pathway. When EHEC is grown on mucins in the presence of *Bacteroides*, virulence genes are upregulated in a FusKR-dependent manner. These results suggest that EHEC uses L-fucose liberated by *Bacteroides* fucosidases to modulate its pathogenicity.¹⁵ This strategy is not limited to L-fucose, as two other enteric pathogens, *Salmonella* Typhimurium and *Clostridium difficile*, have both been shown to use sialic acids that were liberated by sialidases expressed by microbiota.¹³ *S. Typhimurium* shows a significant upregulation of genes involved in the sialic acid catabolism pathway when infecting mice that contain *Bacteroides* species compared to germ-free mice. A similar effect is seen for *C. difficile*

that upregulates its sialic acid catabolism in mice colonized with wildtype *Bacteroides thetaiotaomicron* compared to mice colonized with a sialidase deficient *Bacteroides* mutant.¹³

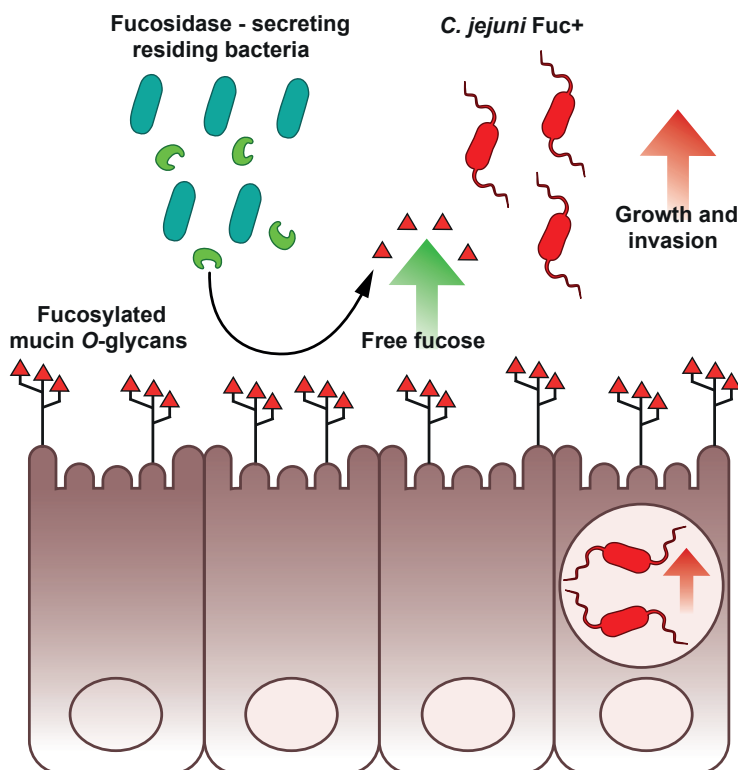


Figure 7. Schematic model of *Bacteroides fragilis* fucosidases enhancing growth and invasion of *Campylobacter jejuni* fuc+ strains. Fucosylated mucin O-glycans (red triangles) are expressed on the apical surface of intestinal epithelial cells. Residing bacteria, such as *Bacteroides fragilis* secrete fucosidases (green), thereby liberating L-fucose from the mucin O-glycans. Our data show that L-fucose and exogenous fucosidases stimulate increased growth and invasion of *C. jejuni* fuc+ strains scavenge and metabolise L-fucose and alter their invasive properties.

Differences in host tropism of *C. jejuni* strains and host mucin composition underscore the importance of choosing the right intestinal epithelial model to investigate virulence of *C. jejuni* fuc+ strains. In a piglet model of human disease, the NCTC 11168 fuc+ strain has been shown to possess a competitive advantage when colonizing the intestinal tract.²² In our studies, we used a confluent monolayer of human intestinal Caco-2 cells and observed that invasion of a *C. jejuni* fuc+ strains was enhanced by fucosidase activity from a co-culture with *B. fragilis*. In chicken, *C. jejuni* colonizes the intestinal tract as

a commensal and hence does not invade. Chicken mucins have an inhibitory effect on *C. jejuni* invasion into epithelial cells and in these animals *fuc+* strains do not have a competitive advantage over other *C. jejuni* strains. However, when chickens were fed additional L-fucose the *C. jejuni fuc+* wildtype strain was more effective in colonizing the intestinal tract compared to a fucose permease knock out strain.^{22,39} Compared to human mucins, chicken mucins contain a large amount of sulfate modifications on their mucin *O*-glycans.⁴⁰ One hypothesis is that the sulfate modifications block the function of exogenous fucosidases.⁴¹ This hypothesis is supported by a mouse model where a decrease in sulfation enhanced intestinal penetrability by pathogens, including *C. jejuni*.⁴² Interestingly, the occurrence of inflammatory bowel disease in humans, which is characterized by decreased mucus barrier function, has been correlated to an altered microbiota with an increased sulfate-reducing bacterial population.⁴³ Therefore, differences in mucin *O*-glycan sulfation and surrounding microbiota could contribute to *C. jejuni fuc+* host tropism. The accessibility of terminal L-fucosides on mucin *O*-glycans for bacterial fucosidases in the presence or absence of specific sulfation patterns is therefore an interesting area for future studies.

C. jejuni is the leading cause of human bacterial gastroenteritis, but our understanding of *C. jejuni* pathogenicity is limited. Our work shows that release of L-fucose by secreted fucosidases from commensal bacteria is an important determinant in *C. jejuni fuc+* growth and invasion (Fig. 6). The novel insight that *C. jejuni fuc+* strains are in part dependent on commensal fucosidases and that these enzymes can be targeted with tailor-made inhibitors provides opportunities for further research. The development of potent and selective bacterial fucosidase probes and inhibitors could lead to novel diagnostic and intervention strategies to target hyperinvasive *C. jejuni fuc+* infections.

Material and Methods

Bacterial strains and media.

Campylobacter jejuni strains NCTC 11168 (clinical strain), 129108 (intestinal isolate from patient with recurrent infections) were used in this study^{44,45}. The complete sequence of the *fuc* operon of *C. jejuni* 108 is deposited at GenBank (CP053854). Routine growth of *C. jejuni* was carried out on 5% saponine-lysed horse blood agar plates (Biotrading) or in HI liquid medium, supplemented with chloramphenicol (20 µg/mL), and/or kanamycin (25 µg/mL) as needed. Cultures were incubated at 37°C in microaerobic incubators with a gas concentration of 80% N₂, 10% CO₂, 5% O₂, 5% H₂. *Bacteroides fragilis* (NTCT9343) and *Bacteroides thetaiotaomicron* (DSMZ 2079) were grown anaerobically at 37°C in basal medium (tryptone yeast extract glucose, TYG)⁴⁶ or Mega medium⁴⁷.

Generation of *C. jejuni* deletion strains.

A *C. jejuni* deletion mutant in the gene *cj0486* was constructed by insertion of a chloramphenicol resistance cassette into the *cj0486* open reading frame. The target

genes were PCR-amplified using the primers YL017_Fw (AGCAAGTTTGAGCATGATAG) and YL018_Rv (TCTTCTAAAGAAGCGCTAGC) with genomic DNA of strains NCTC 11168 and 108 as template. The PCR products were cloned into the pJET1.2 plasmid using the manufacturer's protocol. The deletion plasmids were created by blunt insertion of the *Sma*I-cut chloramphenicol cassette from pAV35 into the *Spe*I-cut pJET1.2-*cj0486*. The knock-out plasmids were transformed into the corresponding *C. jejuni* 11168 and 108 strains by electroporation and transformants were selected on the ability to grow on chloramphenicol containing selective agar plates. *C. jejuni* deletion strains were confirmed by PCR analysis.

L-Fucose growth assays.

DMEM medium (GIBCO) was supplemented with 10 mM L-fucose (Sigma-Aldrich, Netherlands) and filter sterilized (0.2 μ m pore size). Wild type and mutant strains were grown for four days on saponine agar plates supplemented with corresponding antibiotics at 42°C under microaerophilic conditions. Several colonies were picked and plated onto fresh saponine agar plate and grown for 24h at 37°C under microaerophilic conditions. One colony was picked and grown overnight in HI at 37°C. These cultures were used to inoculate the growth medium at an OD₆₀₀ of 0.1. Aliquots of this cell suspension were pipetted into a 96-well plate for use with a synergy HTX plate reader. The plates were placed inside the hypoxic glove box (10% CO₂, 5% O₂, 85% N₂) in a plate reader and incubated at 37°C with moderate, continuous shaking for 20h with OD measurements every 10 minutes. Each growth condition was assessed in triplo and three biological replicates were performed. Statistical analysis was performed using a Student *t* test.

4-Umbelliferon fucopyranoside assay for fucosidase activity.

The enzymatic activity of α -L-fucosidases was assayed at 37°C by incubation with 100 μ M 4-methylumbelliferyl- α -L-fucopyranoside as substrate (Sigma-Aldrich) in 150 mM Mcllvain buffer, pH 4.5. The recombinant α -(1-2,3,4,6)- L-Fucosidase (*Homo Sapiens* fucosidase FucHS; Megazyme) was diluted 1:500 to determine its activity. To determine fucosidase activity in *B. fragilis* or *B. thetaiotaomicron* supernatant fractions, concentrated supernatant (40 times over 10K MWCO spinfilters) of an anaerobic overnight bacterial culture at 37°C of was mixed 1:1 with the pH 4.5 Mcllvain buffer. As a negative control the samples were boiled at 95°C for 10 minutes. The reactions were quenched by adding excess bicarbonate buffer (pH 10), after which fluorescence was measured with a fluorimeter Fluorstar Omega (BMG Labtech) using λ_{ex} 366 nm and λ_{em} 445 nm.

Quantification of fucosidase activity by in-gel fluorescence.

Recombinant α -(1-2,3,4,6)- L-Fucosidase (FucHS; Megazyme) and secreted fucosidases of *B. fragilis* were prepared as described above followed by incubation with 2 μ M of the fluorescent activity-based probe JJB256²⁹ for 30 min at 37°C. The samples were

denatured with Laemmli buffer for 5 min at 95°C and run in the dark on a 10% SDS-PAGE gel. The gels were scanned with an Amersham imager in the CY2 channel. Signal intensity was quantified using ImageJ software.

Mucin growth assays.

10 mg porcine gastric mucin (PGM; Sigma Aldrich) was suspended in 1 mL milliQ and UV-killed 4 times at 100.000 μ Joule in a Stratalink (Stratagene). For fucosidase treatment of mucins with recombinant enzyme, 50 mU α -(1-2,3,4,6)-L-Fucosidase (FucHS; Megazyme;) was added to the 10 mg/mL mucin solution and incubated for 18h at 37°C. To harvest secreted fucosidases of *B. fragillis*, a 6 mL anaerobic overnight culture of *B. fragillis* grown in Mega medium at 37°C was concentrated 60 times using 10K MWCO spinfilters (ThermoFisher Scientific). As a negative control, FNJ (CAS 99212-30-3, Carbosynth) was added to the secreted fraction with a final concentration of 100 μ M to inhibit fucosidase activity. The concentrated secreted fraction was added to the 10 mg/mL mucin aliquot in the ratio 1:8 and incubated for 18h at 37°C. The 10 mg/mL mucin aliquots were diluted in DMEM medium to obtain a final concentration of 1 mg/mL treated or untreated mucins. *C. jejuni* 108 was prepared for growth assays as described above and added to the DMEM medium containing treated or untreated mucins at an OD₆₀₀ of 0.01. Growth of *C. jejuni* was quantified by counting colony-forming units (CFU) on saponin plates incubated at 37°C under microaerophilic conditions for 24h. Statistical analysis was performed using a Student *t* test.

Mammalian cells and culture conditions.

The human gastrointestinal epithelial cell line Caco-2 (ATCC-HTB-37) was routinely cultured in 25 cm² flasks in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal calf serum (FCS) at 37°C in 10% CO₂. For *C. jejuni* gentamicin protection invasion assays, Caco-2 cells were split into 6-well plates and grown for 5 days to form a monolayer before *C. jejuni* infection. For microscopy analysis, cells were cultured on circular glass coverslips in 24-well plates.

Immunofluorescence of fucosylation levels on intestinal epithelial cells.

Caco-2 cells were grown on coverslips as described above and fixed with 4% paraformaldehyde in PBS (Affimetrix) for 30 min at room temperature. Cells were washed twice with Dulbecco's Phosphate Buffered Saline (DPBS, D8537, Sigma Aldrich) before they were permeabilized in binding buffer with 0.1% saponin (Sigma Aldrich) and 0.2% BSA (Sigma Aldrich) in PBS for 30 minutes. Next, coverslips were incubated with the blood group antigen H (O) type 1 mouse monoclonal antibody biotin (Invitrogen 17-206) or UEA-I conjugated with Texas red (EY laboratories T-2201-2) for 1h, followed by 4 washing steps with binding buffer. The coverslips were incubated with streptavidin-488 (Thermo Fisher S11223) 1h at room temperature. Coverslips were washed three times with PBS and once with MilliQ and embedded in Prolonged diamond mounting solution

(Thermo Fisher Scientific). Images were collected on LEICA SPE-II confocal microscope in combination with LEICA LAS AF software.

***C. jejuni* infection assays.**

C. jejuni cultures were grown in HI for 24h at 37°C under microaerophilic conditions and adjusted to OD₆₀₀ of 0.05 in 1 mL DMEM medium. 5 Day-grown Caco-2 cells were washed twice with DMEM without FCS. Bacteria were added to the cells in DMEM without FCS (+/- 10 mM L-fucose; +/- 1 µl FucHS) at MOI= 100 and incubated under microaerophilic conditions at 37°C for 3.5 h. Cells were washed five times with DMEM without FCS and replaced with DMEM without FCS containing 250 µg/mL gentamicin and incubated for 3h at 37 °C to kill extracellular bacteria. Cells were then washed three times with PBS and lysed with 0.5% Triton X-100 in PBS for 5 min at 37°C. Serial dilutions were made and plated on saponine agar plates that contained appropriate antibiotics. Plates were incubated at 37°C under microaerophilic conditions and the number of colony-forming units (CFU) determined. Statistical analysis was performed using a Student *t* test.

***B. fragilis* and *C. jejuni* co-culture infection assay.**

Caco-2 cells and *C. jejuni* 108 were prepared as mentioned above, *B. fragilis* (OD₆₀₀ 1.0 of overnight culture; MOI 10) was added before addition of *C. jejuni*. FNJ (CAS 99212-30-3, Carbosynth) was added to the medium simultaneously with *B. fragilis* with a final concentration of 100 µM, to inhibit fucosidase activity. Invasion of *C. jejuni* was determined as described above.

Acknowledgments

We thank Esther van 't Veld and Richard Wubbolts of the Centre for Cell Imaging (CCI) of the Faculty of Veterinary Medicine for technical support.

References

- 1 M. E. V. Johansson, *PLoS One*, 2012, **7**, e41009.
- 2 M. E. V. Johansson, H. Sjövall and G. C. Hansson, *Nat. Rev. Gastroenterol. Hepatol.*, 2013, **10**, 352–361.
- 3 A. P. Corfield, *Microorganisms*, 2018, **6**, 78.
- 4 M. E. V. Johansson, J. M. Holmén Larsson and G. C. Hansson, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 4659–4665.
- 5 S. Cornick, A. Tawiah and K. Chadee, *Tissue Barriers*, 2015, **3**.
- 6 M. E. V. Johansson, J. K. Gustafsson, J. Holmen-Larsson, K. S. Jabbar, L. Xia, H. Xu, F. K. Ghishan, F. A. Carvalho, A. T. Gewirtz, H. Sjövall and G. C. Hansson, *Gut*, 2014, **63**, 281–291.
- 7 J. A. Ferreyra, K. M. Ng and J. L. Sonnenburg, *Cell. Microbiol.*, 2014, **16**, 993–1003.
- 8 I. Thiele, J. B. Mason, X. Lin, N. Juge, L. E. Tailford, E. H. Crost and D. Kavanaugh, *Front. Genet.* / www.frontiersin.org, 2015, **6**, 81.
- 9 K. S. B. Bergstrom and L. Xia, *Glycobiology*, 2013, **23**, 1026–1037.
- 10 C. Josenhans, J. Muthing, L. Elling, S. Bartfeld and H. Schmidt, *Int. J. Med. Microbiol.*, 2020, **310**, 151392.
- 11 P. B. Eckburg, E. M. Bik, C. N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S. R. Gill, K. E. Nelson and D. A. Relman, *Science (80-.)*, 2005, **308**, 1635–1638.
- 12 L. E. Comstock and D. L. Kasper, *Cell*, 2006, **126**, 847–850.
- 13 K. M. Ng, J. A. Ferreyra, S. K. Higginbottom, J. B. Lynch, P. C. Kashyap, S. Gopinath, N. Naidu, B. Choudhury, B. C. Weimer, D. M. Monack and J. L. Sonnenburg, *Nature*, 2013, **502**, 96–99.
- 14 X. Li, N. M. C. Bleumink-Pluym, Y. M. C. A. Luijckx, R. W. Wubbolts, J. P. M. van Putten and K. Strijbis, *PLoS Pathog.*, 2019, **15**, e1007566.
- 15 A. R. Pacheco, M. M. Curtis, J. M. Ritchie, D. Munera, M. K. Waldor, C. G. Moreira and V. Sperandio, *Nature*, 2012, **492**, 113–117.
- 16 D. Acheson and B. M. Allos, *Clin. Infect. Dis.*, 2001, **32**, 1201–1206.
- 17 J. Silva, D. Leite, M. Fernandes, C. Mena, P. A. Gibbs and P. Teixeira, *Front. Microbiol.*, 2011, **2**.
- 18 N. O. Kaakoush, N. Castaño-Rodríguez, H. M. Mitchell and S. M. Man, *Clin. Microbiol. Rev.*, 2015, **28**, 687–720.
- 19 M. Stahl, J. Butcher and A. Stintzi, *Front. Cell. Infect. Microbiol.*, 2012, **2**, 5.
- 20 D. Hofreuter, *Front. Cell. Infect. Microbiol.*, DOI:10.3389/fcimb.2014.00137.
- 21 W. T. Muraoka and Q. Zhang, *J. Bacteriol.*, 2011, **193**, 1065–1075.
- 22 M. Stahl, L. M. Friis, H. Nothaft, X. Liu, J. Li, C. M. Szymanski and A. Stintzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 7194–7199.
- 23 J. J. J. van der Hooft, W. Alghafari, E. Watson, P. Everest, F. R. Morton, K. E. V. Burgess and D. G. E. Smith, *Metabolomics*, DOI:10.1007/s11306-018-1438-5.
- 24 J. M. Garber, H. Nothaft, B. Pluvinae, M. Stahl, X. Bian, S. Porfirio, A. Enriquez, J. Butcher, H. Huang, J. Glushka, E. Line, J. A. Gerlt, P. Azadi, A. Stintzi, A. B. Boraston and C. M. Szymanski, *Commun. Biol.*, 2020, **3**, 1–11.
- 25 L. N. Nielsen, T. A. Luijckx, C. S. Vegge, C. K. Johnsen, P. Nuijten, B. W. Wren, H. Ingmer and K. A. Krogfelt, *Clin. Vaccine Immunol.*, 2012, **19**, 113–119.

- 26 C. Fearnley, G. Manning, M. Bagnall, M. A. Javed, T. M. Wassenaar, D. G. Newell, M. Afzal Javed, T. M. Wassenaar, D. G. Newell, C. G. Diane Newell, M. A. Javed, T. M. Wassenaar and D. G. Newell, *J. Med. Microbiol.*, 2008, **57**, 570–580.
- 27 M. A. Javed, A. J. Grant, M. C. Bagnall, D. J. Maskell, D. G. Newell and G. Manning, *Microbiology*, 2010, **156**, 1134–1143.
- 28 W. Elhenawy, M. O. Debelyy and M. F. Feldman, *MBio*, DOI:10.1128/mBio.00909-14.
- 29 J. Jiang, W. W. Kallemijn, D. W. Wright, A. M. C. H. C. H. Van Den Nieuwendijk, V. C. Rohde, E. C. Folch, H. Van Den Elst, B. I. Florea, S. Scheij, W. E. Donker-Koopman, M. Verhoek, N. Li, M. Schürmann, D. Mink, R. G. Boot, J. D. C. C. Codée, G. A. Van Der Marel, G. J. Davies, J. M. F. G. F. G. Aerts, H. S. Overkleef, V. Coco Rohde, E. Colomina Folch, H. Van Den Elst, B. I. Florea, S. Scheij, W. E. Donker-Koopman, M. Verhoek, N. Li, M. Schürmann, D. Mink, R. G. Boot, J. D. C. C. Codée, G. A. van den Marel, G. J. Davies, J. M. F. G. F. G. Aerts and H. S. Overkleef, *Chem. Sci.*, 2015, **6**, 2782–2789.
- 30 G. M. Ruiz-Palacios, L. E. Cervantes, P. Ramos, B. Chavez-Munguia and D. S. Newburg, *J. Biol. Chem.*, 2003, **278**, 14112–20.
- 31 J. Amano and M. Oshima, *J. Biol. Chem.*, 1999, **274**, 21209–21216.
- 32 D. C. Mills, O. Gundogdu, A. Elmi, M. Bajaj-Elliott, P. W. Taylor, B. W. Wren and N. Dorrell, *Infect. Immun.*, 2012, **80**, 1690–1698.
- 33 L. V Hooper, J. Xu, P. G. Falk, T. Midtvedt and J. I. Gordon, 1999, **96**, 9833–9838.
- 34 A. Lee, J. L. O'Rourke, P. J. Barrington and T. J. Trust, *Infect. Immun.*, 1986, **51**, 536–546.
- 35 X. Bian, J. M. Garber, K. K. Cooper, S. Huynh, J. Jones, M. K. Mills, D. Rafala, D. Nasrin, K. L. Kotloff, C. T. Parker, S. M. Tennant, W. G. Miller and C. M. Szymanski, *mSphere*, DOI:10.1128/msphere.00735-19.
- 36 J. Dicksved, P. Ellström, L. Engstrand and H. Rautelin, *MBio*, DOI:10.1128/mBio.01212-14.
- 37 J. L. Sonnenburg, J. Xu, D. D. Leip, C. H. Chen, B. P. Westover, J. Weatherford, J. D. Buhler and J. I. Gordon, *Science (80-.)*, 2005, **307**, 1955–1959.
- 38 B. F. Cravatt, A. T. Wright and J. W. Kozarich, *Annu. Rev. Biochem.*, 2008, **77**, 383–414.
- 39 C. M. Byrne, M. Clyne and B. Bourke, *Microbiology*, 2007, **153**, 561–569.
- 40 W. B. Struwe, R. Gough, M. E. Gallagher, D. T. Kenny, S. D. Carrington, N. G. Karlsson and P. M. Rudd, *Mol. Cell. Proteomics*, 2015, **14**, 1464–1477.
- 41 A. M. Robertson and D. P. Wright, *Can. J. Gastroenterol.*, 1997, **11**, 361–366.
- 42 P. A. Dawson, S. Huxley, B. Gardiner, T. Tran, J. L. McAuley, S. Grimmond, M. A. McGuckin and D. Markovich, *Gut*, 2009, **58**, 910–919.
- 43 N. Ijssennagger, R. van der Meer and S. W. C. van Mil, *Trends Mol. Med.*, 2016, **22**, 190–199.
- 44 H. P. Endtz, B. A. J. J. Giesendorf, A. Van Belkum, S. J. M. M. Lauwers, W. H. Jansen and W. G. V. V Quint, *Res. Microbiol.*, 1993, **144**, 703–708.
- 45 M. B. Skirrow, *BRITISH MEDICAL JOURNAL Campylobacter enteritis: a 'new' disease*, 1977.
- 46 M. K. Bacic and C. J. Smith, *Curr. Protoc. Microbiol.*, 2008, **0 13**.
- 47 M. Wu, N. P. McNulty, D. A. Rodionov, M. S. Khoroshkin, N. W. Griffin, J. Cheng, P. Latreille, R. A. Kerstetter, N. Terrapon, B. Henrissat, A. L. Osterman and J. I. Gordon, *Science (80-.)*, DOI:10.1126/science.aac5992.

Chapter 3

***O*-glycomic signatures of sodium butyrate-stimulated enterocyte differentiation**

Yvette M.C.A. Luijkx[#], Katarina Madunić[#], Guinevere S.M. Lageveen Kammeijer, Oleg A. Mayborodaa, Tom Wennekes, Karin Strijbis, Manfred Wuhrer

[#] These authors contributed equally to this work

Abstract

The gut microbiota residing in the human gastrointestinal tract provide essential health benefits to its host, particularly by excreting bacterial metabolites. The well-known bacterial metabolite butyrate has shown to have a beneficial effect on the intestinal homeostasis. It is the preferred energy source of intestinal epithelial cells, capable of inducing differentiation, and has been correlated with upregulating expression of highly *O*-glycosylated mucin proteins. Mucins and their glycosylation have previously been established as key players in the complex processes of differentiation. We therefore hypothesized that GI tract epithelial cell differentiation, and thereby homeostasis, is in part controlled by a complex underlying network of bacterial metabolites, mucin expression and *O*-glycosylation. Additionally, mucins and their *O*-glycosylation have also been identified as key players in carcinogenesis, and butyrate was shown to have the capacity to reduce colorectal carcinoma potential. The role of specific mucins and their *O*-glycan structures in healthy differentiation, homeostasis, and carcinogenesis however remains to be established. The aim of the study presented in this chapter is to identify specific mucin *O*-glycan structures that mark butyrate-induced epithelial differentiation of the intestinal cell line Caco-2. An in-depth *O*-glycan analysis of spontaneously differentiated and butyrate-differentiated Caco-2 cells revealed significant differences in *O*-glycosylation. Fully differentiated butyrate-stimulated cells are characterized by a higher expression of the Sda antigen and upregulation of disialylated structures compared to the spontaneous differentiated samples. This unbiased in-depth analysis of the Caco-2 *O*-glycome provides evidence for a specific butyrate-dependent *O*-glycan signature. This insight and overall approach paves the way for future endeavors to study the dynamic *O*-glycosylation patterns in the gut, either produced via cellular biosynthesis or through the action of bacterial glycosidases, and the functional role of these patterns in homeostasis and dysbiosis at the gut-microbiota interface.

Introduction

The human gut microbiota is a complex ecology of a variety of different microorganisms. Among viruses, archaea, and eukarya, bacteria are the most abundant inhabitant of the human gastrointestinal (GI) tract. By co-evolution with the host, a symbiotic relationship formed between the GI tract and its bacteria.^{1,2} The maintenance of homeostasis in the GI tract depends in part on the complex process of epithelial cell differentiation.³ Bacterial metabolites have also been identified as playing a functional role in the maintenance of this homeostasis.⁴ The short chain fatty acid, butyrate, is a bacterial metabolite produced by bacterial fermentation of dietary fibers and is known to have a beneficial effect on the intestinal homeostasis. Butyrate was also shown to have anti-tumor properties. Physiological concentrations of butyrate inhibit proliferation, induce differentiation and increase the rate of apoptosis for a number of tumor cell types *in vivo* and *in vitro*.^{5–8} Furthermore, butyrate is the main energy source of intestinal epithelial cells and capable of upregulating the gene expression of both secreted and membrane-linked mucins.^{9–13} Noticeably, bacterial short chain fatty acids such as butyrate, are suspected of also extending their effect outside of the gut as they have been implicated as possible mediators of phenomena observed in the gut-brain axis.¹⁴

The mucus layer covers the epithelial cells to prevent them from being in direct contact with the microbiota. The barrier function of the mucosal layer is largely maintained by the gel-forming mucins, which are complex *O*-glycosylated proteins secreted by the epithelial cells.¹⁵ Furthermore, epithelial cells express membrane-linked mucins that have both a barrier and a signalling function.¹⁶ The intracellular tail of the membrane-linked mucins can be phosphorylated and connects to signalling pathways that regulate cell-cell interactions, differentiation, and apoptosis.¹⁶ Additionally, the mucus layer and especially its *O*-glycans is an important nutrient source for the surrounding microbiota and thereby contributes to bacterial colonization in the human gut. Mucins are often composed of tandem repeat sequences rich in serine, proline, and threonine residues that are highly *O*-glycosylated. More than 80% of the mucin dry weight is made up of *O*-glycans and *O*-glycosylation is thereby the main modification that affects mucin properties and function.¹⁷ A wide variety of oligosaccharide structures can be attached to the mucins and the composition of these structures can vary within cell types and specific niches due to differential expression of glycosidases and glycosyltransferases.¹⁸

Furthermore, membrane-linked mucins are seen as important oncogenic proteins as their intracellular domain and their glycosylated extracellular domain link to pathways involved in cell differentiation and apoptosis.^{16,19,20} Noticeable, altered glycosylation can be correlated back to the bacterial metabolite butyrate levels, as this short chain fatty acid regulates specific genes involved in the glycosylation pathways including the lectin galectin-1 and glycosyltransferase β -galactoside- α -2,6-sialyltransferase.^{21–24} It is known that butyrate regulates gene expression through its function as an inhibitor of the

chromatin remodeling activity of histone deacetylases (HDACs).^{25,26} It could therefore be hypothesized that in the GI tract differentiation is controlled by a complex underlying network of bacterial metabolites, mucin expression and *O*-glycosylation.

The intestinal cell line, Caco-2 (*Cancer coli-2*), that is investigated in this study was derived from a human colorectal carcinoma (CRC) in 1977 and has since become a well-established model to study cellular differentiation. The cell line can differentiate spontaneously or by exposure to butyrate into polarized cells with morphological and biochemical features of the mature colonic epithelium.²⁷ Spontaneous differentiation of Caco-2 cells has been studied previously, including transcriptomics, proteomics and glycomics analysis. The associated changes in *O*-glycosylation correlated to butyrate differentiation have yet to be characterized and insight into this is crucial for a better understanding of the role of glycans in gut homeostasis.^{28–35} The aim of the study presented in this chapter was to identify specific *O*-glycan structures that define butyrate-induced epithelial differentiation to gain more insight in *O*-glycan signatures of healthy differentiation and colorectal cancer (CRC)-related changes in the gut epithelium.

Results

Cell differentiation assay

O-glycosylation of Caco-2 cells was analyzed at different time points during the process of spontaneous enterocyte-like differentiation and sodium butyrate (NaBu) induced differentiation. Cells were proliferated to confluency, which is marked as day 0, after which cells accelerate their differentiation. Differentiation of Caco-2 cells was evaluated by a colorimetric assay measuring alkaline phosphatase activity (AP), which is a marker for late stage differentiation indicative for the presence of an established brush border.³⁶ AP activity levels of both spontaneous differentiation and induced differentiation showed a continuing increase. At days 5 and 7, a significantly higher level of AP activity for the induced differentiation was observed compared to the spontaneous differentiation (Figure 1A).

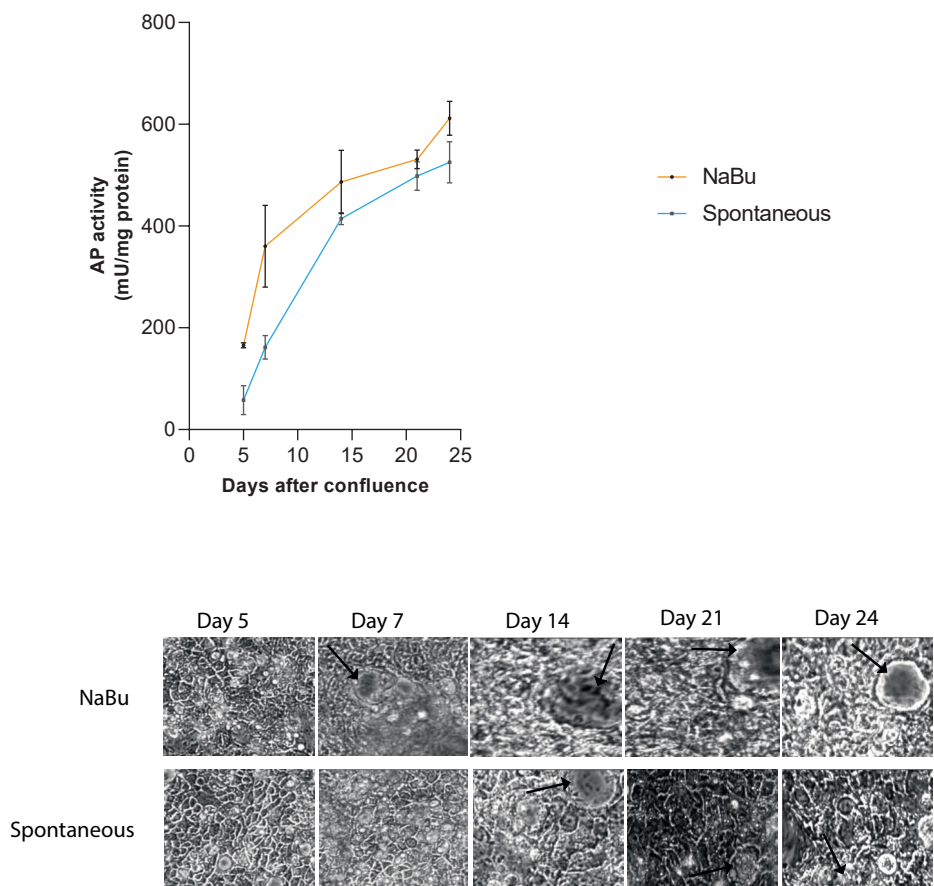


Figure 1 | Biochemical and morphological validation of Caco-2 differentiation. **a)** Alkaline phosphatase (AP) activity during Caco-2 growth. Stimulated by 2 mM sodium butyrate (NaBu, orange line) or spontaneous differentiation (blue line). Standard deviation is indicated by error bars ($n = 3$). **b)** Phase contrast images of Caco-2 cells on culture dishes at different stages of growth. Arrow; dome-like structure. Original magnification; 40x.

In addition to AP activity levels as a validation for differentiation, phase images were taken of the apical surface to observe differentiation induced “dome” formation. Caco-2 cells imaged at different growth times started displaying dome formation at day 7 for the NaBu induced cells, whereas the spontaneously differentiated cells showed dome formation much later at day 14 (Figure 1B). After having established that Caco-2 cells are differentiated our next aim was to assess the effect of butyrate on the *O*-glycosylation profile in these Caco-2 cells.

Glycomic analysis

The analysis of *O*-glycans can be challenging in comparison to *N*-glycans, due to the lack of specific enzymes that can release the intact *O*-glycans from glycosylated proteins. In this study, we used a previously established protocol that allowed the chemical release of *O*-glycans from cell lysates via reductive β -elimination in 96 well plate format.³⁴ *O*-glycans of the differentiating Caco-2 cells were analyzed at days 5, 7, 14, 21 and 24 post confluency. To further validate the robustness of this approach for our samples, full technical triplicates of the cell lysate containing 0.5 million cells, from the same biological replicate (5 days, spontaneous differentiation), were processed and analyzed independently by PGC nano-LC–ESI-MS/MS. The mean relative area of the 11 most abundant glycan species and corresponding standard deviations are shown in Figure S1a. This data set indicates a low technical variability of the workflow. Parallel to this, the system performance was assessed by releasing *O*-glycans from 10 μ g of bovine submaxillary mucin (BSM) standard and measured each day across 5 days, as illustrated in Figure S1b.

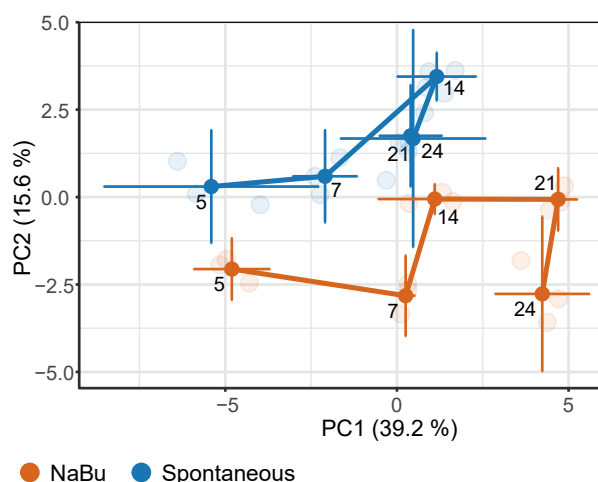


Figure 2 | The geometric trajectory visualization of the PCA scores. PCA model based on relative abundance (%) of different glycans shows a separation between the sodium butyrate (NaBu) stimulated group (orange) and spontaneous differentiation (blue). The separation between different timepoints in the two groups is illustrated as a trajectory. The top two principal components (PC) explain 54.8% of the variation within the data. The PCA scores from different biological replicates (faded color) were averaged to create the trajectory.

To explore specific *O*-glycan variations that correlate to the differentiation of NaBu-stimulated samples, a principal component analysis (PCA) was used. The first two principal components (PC) explain more than 54.8% of the total variance. The model illustrates relatively close clustering of biological replicates, although substantial differences are observed within replicates for timepoints 24 days (NaBu) and 5 days (spontaneous), see Figure S2a. Furthermore, there is a clear separation between the

NaBu-stimulated samples and spontaneous differentiated samples along the PC2 (15.6 %). Different timepoints of differentiation are separated along the PC1 (39.2%). A geometric trajectory (Figure 2) illustrates the separation between the two groups (NaBu-stimulated and spontaneous differentiation) along PC2, and between timepoints along PC1. To decipher which glycan features drove the separation in the PCA model, we explored the variables in the PCA loading plot (Figure S2b). Samples clustering in the upper left corner of the plot show a higher relative abundance of glycans containing α L-fucose residue, whereas the samples clustering on the lower right part of the plot show a relative higher abundance of sialylated glycan species.

Further analysis of the scores plot (Figure S2a) showed that for days 5 and 7 both groups are different from one another as well as from days 14, 21 and 24, since the representations do not overlap. This indicates that the Caco-2 *O*-glycome undergoes progressive changes when cells are grown from confluency to 7 days post-confluence, at which point these cells are not fully differentiated yet. Upon further differentiation a distinct difference between the NaBu-stimulated group trajectory and the spontaneous differentiation can be observed. For the spontaneous differentiated Caco-2 cells the representations of days 14, 21 and 24 are close to one another. This indicates that the glycome appears to stabilize in the later differentiation phases. However, the NaBu-stimulated cells are different from each other between days 14, 21 and 24, which indicates, the NaBu-stimulated cells still undergo progressive changes in their *O*-glycome in the late differentiation phases. To support this visual interpretation and identify the glycosylation signatures changing with differentiation, two-way ANOVA analysis of variance was performed on relative abundances of individual *O*-glycans. Specific *O*-glycans that show statistically significant changes, either with time or difference between groups (NaBu-stimulated vs. spontaneous differentiation), are summarized in Figure S3.

O-glycans that follow significant trends with differentiation were selected for visualizations (Figure 3 and 4). Figure 3 shows *O*-glycans that are significantly downregulated with differentiation and are carriers of terminal blood group antigens (H2N2F1b, H2N2F1d, H2N2F2b, H2N2F1S1b, H2N2F1S1c), as well as sulphation (H2N2S1Su1).

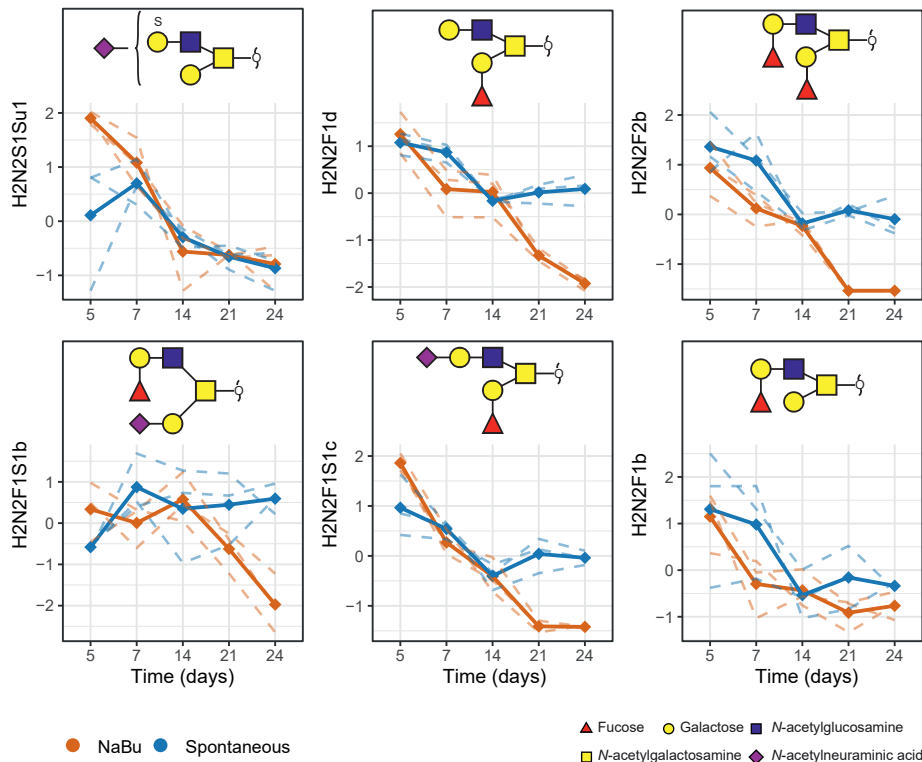


Figure 3 | Differentiation induces significant downregulation of Caco-2 O-glycans. Spaghetti plots of the glycan abundances that show significant difference between timepoints selected from the analysis of variance (ANOVA). The dashed lines represent the scaled z-scores of the measured values of each biological replicate, whilst the continuous lines represent the z-scores of the mean values per biological replicate.

The O-glycans that are significantly upregulated with differentiation as well as those that show a difference between the groups (NaBu-stimulated and spontaneous differentiation) are depicted in Figure 4. All of the illustrated O-glycans carry terminal N-acetylneuraminic acids (Neu5Ac) and no fucosylation (H2N2S2b, H2N2S2a, H3N3S2b, H1N1S1a, H1N2S2, H1N1S3). Notably, the Sda antigen present on O-glycans with composition H1N2S2 shows downregulation with differentiation. However, the Sda antigen (GalNAc β 1-4(Neu5Ac α 2-3)Gal β -) is significantly higher in the NaBu-stimulated group compared to the spontaneous differentiation. H1N1S1 corresponds to an O-glycan carrying an α 2,6-linked Neu5Ac at the innermost N-acetylgalactosamine (GalNAc) and shows significant upregulation with differentiation. Furthermore, one specific O-glycan (H1N1S3) shows an upregulation that is only visible in the NaBu-stimulated group. By MS/MS fragmentation we could confirm that this structure carries two sialic acids linked together, probably via α 2,8-linkage (Figure S4).³⁷ Corresponding with the previous observed trajectory in the PCA models, the plots in Figure 3 and 4 both show

a stabilization in the *O*-glycome profile of spontaneous differentiated cells after day 14. Whereas the NaBu-stimulated cells continue to progress after day 14.

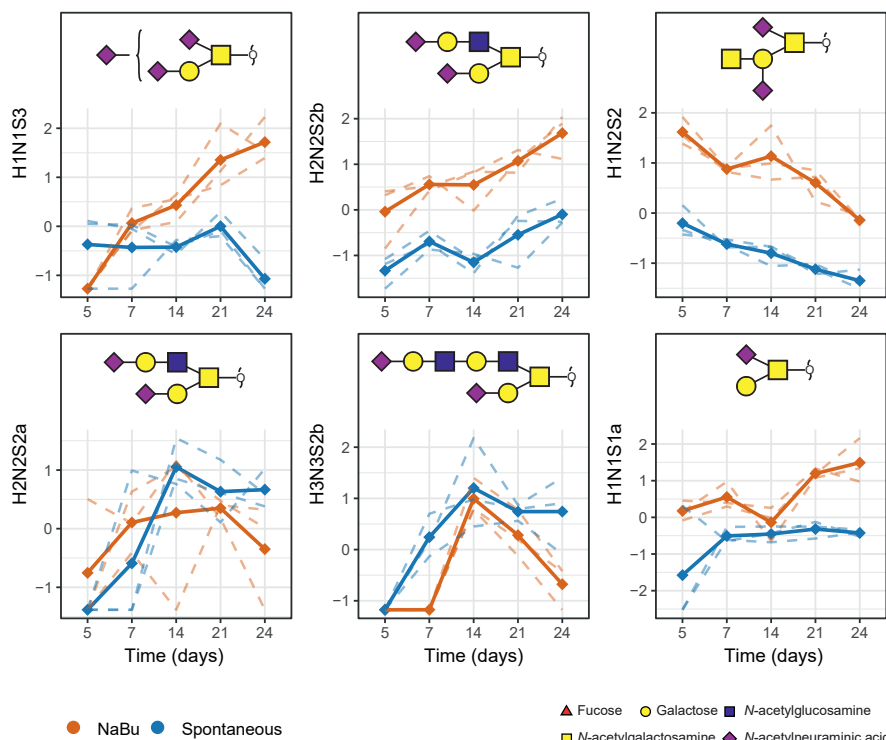


Figure 4 | Differentiation induces significant upregulation of Caco-2 *O*-glycans as well as some NaBu specific changes. Spaghetti plots of the glycan abundances that show significant difference between groups selected from the analysis of variance (ANOVA). The dashed lines represent the scaled z-scores of the measured values of each biological replicate, whilst the continuous lines represent the z-scores of the mean values per biological replicate.

Structural identification of glycan species

Due to reports from the literature that provide contradictory *O*-glycan assignments, we made additional efforts towards the identification of specific *O*-glycan structures in our study by comparison to the *O*-glycan repertoire of BSM that has been very well characterized by both NMR and tandem mass spectrometry.^{38–40}

Two isomers with composition H2N2F1 at m/z 895.34 are observed in our Caco-2 samples (1 and 2a) (Figure S5). The major isomer of expressed in Caco-2 cell line (2a), matched with the major isomer in BSM standard (2b) by retention time and MS2 which allowed us to confirm that the major *O*-glycan isomer 2a in Caco-2 sample carried the terminal blood group H antigen.³⁹ The MS2 spectra of this isomer shows the presence of a high abundant Z-ion (m/z 569.20 (H1N2) indicating the occupancy of the 6'-antenna

of the core 2 *O*-glycan. On the other hand, isomer 1 is expressed solely by Caco-2 cells, and the MS/MS spectra revealed the presence of terminal fucose linked to a type 2 *N*-acetylglucosamine on the 6' antenna. Indicated by the characteristic cross-ring fragment of the β 1,4-linked galactose (m/z 409.02). However, since this glycan is not expressed in the BSM standard, it is not possible to determine the linkage of the terminal fucose residue. Previous studies have reported an α 1,2-linked terminal L-fucose, as part of blood group antigen H.⁴¹ However, A recent study has reported the presence of an unconventional terminal α 1,6-linked fucose.³³

Additionally, two isomers with composition H₂N₂F₂ at m/z 1041.40 are observed in our Caco-2 samples (1 and 2a) Figure S6. Isomer 2a matched by retention time and MS/MS spectra with the major isomer in BSM standard (2b), indicating a blood group antigen H related terminal α 1,2-fucose.³³ Isomer 1 could not be identified by MS2 nor provide insightful differences in the fragmentation patterns due to low abundance. Interestingly, a recent study reported presence of two α 1,6- linked terminal fucosides carried by a major isomer with composition H₂N₂F₂ in Caco-2.³³

Figure S6 shows the structural identification of an *O*-glycan with composition H₁N₁F₁S₁ at m/z 821.30 carrying blood group antigen type 3 indicated by the fragments at m/z 530.23 and 512.14, where m/z 495.25 indicates an α 2,6-linked sialic acid on the innermost GalNAc, which is in contrast to a previously reported annotation where fucose is linked to the innermost GalNAc.³³

Structural elucidation of a glycan with composition H₁N₁S₃ at m/z 1257.44 is shown in Figure S7. Two sialic acids seem to be linked to each other, indicated by the presence of a characteristic fragment ion at m/z 581.18. Although, there is not sufficient evidence to support the location of the epitope, it has been previously described in literature that this occurs on the α 2,6-linked sialic acid linked to the innermost GalNAc.³⁷

Overall, we observed a downregulation of terminal blood group H antigen expression and an upregulation of both galactose α 2,3-sialylation and core GalNAc α 2,6-sialylation with differentiation.

Discussion

In this study we observed a clear *O*-glycome profile correlated to differentiation. Differentiation of the Caco-2 cells was accompanied by the elevation of galactose α 2,3-sialylation, upregulation of GalNAc α -2,6 sialylation, and a downregulation of the terminal blood group H antigen. The first two observations have previously been correlated to differentiation by the expression of their corresponding glycosyltransferases.^{42–44} However, the published studies on correlation of the terminal blood group H antigen to differentiation are contradictory.^{45–47} It was previously postulated that Caco-2 cells

are a heterogeneous cell population and considerable differences between laboratories has been seen before.⁴⁸

Additionally, we focused on distinct differences between spontaneous and NaBu-stimulated differentiation. The PCA plots and two-way ANOVA on the relative abundances of individual *O*-glycans, indicated that the Caco-2 *O*-glycome for spontaneous differentiation stabilizes after day 14. Which is in agreement with an proteomics analysis previously published by Stierum and coworkers on the spontaneous differentiation of Caco-2 cells.²⁸ The *O*-glycome for the NaBu-stimulated group behaves significantly different and does not stabilize after day 14. We hereafter sought to discuss some of the glycan structures differentially expressed in the butyrate stimulated cells.

The *O*-glycan carrying a Sda antigen shows a downregulation over time for both groups, correlating to differentiation, but a general higher expression for the NaBu-stimulated cells compared to the spontaneous differentiated cells. Most studies report a downregulation of the Sda antigen correlated to differentiation in CRC cell lines, which corresponds to the general downregulation we observed over time.^{35,49} However, none of these studies investigated the level of Sda antigen expression in correlation to NaBu stimulation. It has previously been observed that the expression of the Sda antigen and the B4GALNT2 gene are regulated by promoter DNA methylation.⁵⁰ It was reported that in colon cancer cell lines B4GALNT2 promoter methylation was associated with very low levels of mRNA.⁵⁰ Recent literature reported the capability of butyrate to trigger DNA demethylation in CRC cell lines.⁵¹ Therefore, something similar could be envisioned that explains the upregulation of the Sda antigen between the groups (NaBu and spontaneous) and is in agreement with the reported anti-carcinogenic effect of butyrate.

Another interesting *O*-glycan that shows an upregulation in the NaBu-stimulated samples compared to the spontaneous differentiated samples (after day 14), is the H1N1S3 *O*-glycan. MS/MS fragmentation revealed that this *O*-glycan carries a di-sialic acid. Unlike the other monosaccharides, sialic acids are capable of forming polymers like a di-sialic acid (DiSia) and poly-sialic acid (PolySia).⁵² PolySia has been relatively well-studied, whereas less focus went out to DiSia glycoproteins.^{53,54} PolySia of a cell adhesion molecule has previously been linked to promoting metastasis in CRC.⁵⁵ However, it was postulated that the biological function of DiSia is clearly distinct from polySia, but details remain unknown.⁵² Therefore, this DiSia would be an interesting structure for future studies in correlation to tumorigenic properties.

Noticeable, in the spontaneous differentiated cells the expression of terminal blood group antigens carrying an α -1,2-L-fucoside stabilizes at day 14. Whereas most of these α -1,2-L-fucoside carrying antigens in the butyrate-stimulated cells continue decreasing even after day 14. L-Fucoside moieties on glycans are increasingly recognized as critical

for cell-cell interactions and signaling processes.⁵⁶ As butyrate has previously been linked to exert anti-tumor properties it would be interesting to further investigate the implications of this butyrate-induced fucose downregulation on tumor progression.

Conclusions

Differentiated Caco-2 cells are an important and widely used model cell line. Therefore, it is important to note that attention should be given when using these Caco-2 cell line as changes in glycosylation expression patterns are both growth- and bacterial metabolite-dependent. To conclude, this untargeted in-depth screening of the Caco-2 cell line identified specific *O*-glycan structures that mark butyrate-induced epithelial differentiation. Some interesting structures were identified that will need further studies to identify their role in maintaining homeostasis or CRC-related changes in the epithelium.

When correlating the here obtained data to *in vitro* experiments it should be taken into account that this study focused on the transmembrane mucins, thereby leaving out the possibly differential expression of *O*-glycan on the abundant secreted mucins. It would be very interesting to expand the use of this technique to study *O*-glycosylation of the secretome. In the future this might allow us to better study the constantly changing *O*-glycosylation patterns at the gut-microbiota interface as correlated to bacterial hydrolase activities and its relevance to maintaining homeostasis or role in dysbiosis.

Experimental

General methods and materials

Ammonium bicarbonate (ABC), Dowex cation-exchange resin (50W-X8), trifluoroacetic acid (TFA) Dulbecco's Phosphate Buffered Saline (DPBS), Hydrochloric acid (HCl), DL-dithiothreitol (DTT), sodium chloride (NaCl), and sodium borohydride (NaBH₄) were obtained from Sigma-Aldrich (Steinheim, Germany). 8 M guanidine hydrochloride (GuHCl), Dulbecco's modified Eagle's medium (DMEM), 0.25% trypsin/EDTA, and FCS were purchased from Thermo Fisher Scientific. Peptide N-glycosidase F (PNGase F) and complete EDTA free protease inhibitor cocktail tablets were purchased from Roche Diagnostics (Mannheim, Germany). Potassium hydroxide (KOH) was purchased from Honeywell Fluka. Solid phase extraction (SPE) bulk sorbent Carbograp was obtained from Grace Discovery sciences (Columbia, USA). HPLC SupraGradient acetonitrile (ACN) was obtained from Biosolve (Valkenswaard, The Netherlands). Ethanol (Reag. Ph. Eur) and Mucin from bovine submaxillary glands, type I-S were purchased from Merck (Darmstadt, Germany). MultiScreen HTS 96 multiwell plates (hydrophobic Immobilon-P PVDF membrane) and 96-well PP Microplate were obtained from Millipore (Amsterdam, The Netherlands). 96-well PP filter plate was purchased from Orochem Technologies (Naperville, IL, USA).

Cell culture. Human colorectal adenocarcinoma Caco-2 cells were grown in Dulbecco's Modified Eagle medium (Gibco, Thermo Fisher Scientific) with 10% fetal calf serum. Cells were subcultured at 80% confluency and maintained at 37 °C in a humidified incubator with 5% CO₂. At day 0, when cells reached full confluency, 2 mM NaBu was added to the appropriate cells. On days 5, 7, 14, 21 and 24 cells were collected in biological triplicates. For harvesting of the cells, medium was removed and adherent cells were washed twice with DPBS and trypsinized using 0.25% trypsin- 1mM EDTA. To stop the trypsin activity, medium (without FCS) in a ratio of 2:5 (trypsin:medium; v/v) was added and cells were pelleted at 300 x g for 5 minutes. Cells were resuspended in DPBS and counted and aliquoted to ~2.0 x 10⁶ cells per replicate and washed twice with 1 mL DPBS for 3 minutes at 100 x g. The supernatant was removed and cell pellets stored at -20 °C until further use.

Phase-contrast microscopy. Phase images were acquired using phase contrast microscopy with 40 times objective. Images were acquired through an CMEX 5000 Microscope camera.

Alkaline phosphatase activity. Alkaline phosphatase (ALP) activity was measured in cell homogenates using a colorimetric assay according to the manufacturer's instructions (Abcam). The cells were harvested with 0.25% trypsin- 1mM EDTA, washed with PBS and subjected to ultrasonification in the supplied assay buffer. The obtained cell homogenates were stored at -80°C until assayed. Cells lysed at different growth days were incubated with 1.7 mM *p*-nitrophenylphosphate (*p*-NPP) for 60 min at 25 °C. Absorbance readings were taken at 405 nm with a microplate reader (Fluorimeter Fluorstar Omega, BMG Labtech). Results are expressed as milli units/mg protein. One unit is defined as the activity that hydrolyzes 1 micromole of *p*-NPP/min at 25 °C.

Cell lysis and O-glycan release. Three biological replicates from each timepoint were analyzed. Frozen cell pellets containing ~2 million cells were resuspended in 100 µL of lysis buffer containing Tris HCl, EDTA, NaCl and protease inhibitor cocktail. The cells were lysed using Branson sonicator rod at 1.5, and 25 µL of the suspension containing ~0.5 million cells were loaded onto the preconditioned PVDF membrane plate wells. 10 µg of bovine submaxillary mucin was blotted in 3 different wells of the same PVDF membrane plate. The denaturation, *N*- and *O*-glycan release were performed as described previously.^{34,57} Briefly, the proteins were denatured on membrane using guanidine hydrochloride and DTT at 60 °C. Upon removal of denaturing agents, the *N*-glycans were released by PNGase F digestion overnight at 37 °C, and recovered in MQ water. Upon recovery of *N*-glycans, the *O*-glycans were released from the same wells by reductive beta elimination, using sodium borohydride in potassium hydroxide incubating at 50 °C for 16 h. Sample desalting by cation exchange resin Dowex 50 W X8 and graphitized carbon solid phase extraction were performed in self-packed 96-well filter plates. The samples were dried after cleaning and stored at -20 °C until analysis.

PGC-LC-MS/MS analysis. The *O*-glycan samples were then reconstituted in 20 μ L of MQ water, and 2 μ L were injected for analysis. Analysis was performed using a PGC nano-LC Ultimate 3000 UHPLC system (Dionex/Thermo, Sunnyvale, CA) coupled to an amaZon ETD speed ion trap (Bruker Daltonics, Bremen, Germany). The samples were loaded using 100% buffer A (10 mM ammonium bicarbonate) at a loading flow of 6 μ L/min on a custom-made trap column (size 30 \times 0.32 mm) packed with 5 μ m particle size PGC stationary phase from Hypercarb PGC analytical column (size 100 \times 4.6 mm, 5 μ m particle size, Thermo Fisher Scientific, Waltham, MA). Afterwards, the *O*-glycans were separated at a 0.6 μ L/min flow rate on a custom-made PGC column (100 mm \times 100 μ m, 3 μ m particle size obtained from Thermo Fisher Scientific) by applying a linear gradient from 1% to 50% buffer B (MeCN, 10 mM ammonium bicarbonate) over 73 min. During the procedures a constant column temperature of 45°C was maintained. To continue, the LC system was coupled to an amaZon ETD speed ESI ion trap MS using the CaptiveSpray™ source (Bruker Daltonics) with an applied capillary voltage of 1000 V in negative-ionization mode. The flow rate was set to 3 L/min and the drying gas (N₂) temperature was set at 280°C. The nebulizer gas pressure was kept at 3 psi. The nanoBooster™ bottle (Bruker Daltonics) was filled with methanol, as a dopant solvent. MS spectra were acquired in enhanced mode within a mass to charge ratio (*m/z*) range of 380–1850. The maximum acquisition time was set to 200 ms mass, the target of smart parameter setting was set to *m/z* 900, and the ion charge control (ICC) to 40,000. MS/MS spectra were generated by collision-induced dissociation on the three most abundant precursors, applying an isolation width of 3 Thomson. Additionally, ICC was set to 150 000 and the fragmentation cut-off was set to 27% with a 100% fragmentation amplitude using the Enhanced SmartFrag option (30–120% in 32 ms). To integrate area under the curve (AUC) for each individual glycan isomer, extracted ion chromatograms of the first three isotopes were used in Compass DataAnalysis software v.5.0. Peaks were manually picked and integrated. Total area normalization was employed for relative quantification of *O*-glycan species. Identification of *O*-glycan species was performed by comparison with PGC retention time, MS/MS spectra, and a BSM standard.

Statistical analysis. Data analysis and visualization was performed using in-house developed “R” scripts. To enable use of principal component analysis, imputation of minimum positive number (0.0001) was performed. Differences between groups were tested using two way ANOVA analysis of variance.

Acknowledgements

We thank Katarina Madunic, Guinevere Kammeijer, and Oleg Mayborodaa for the MS analysis and interpretation.

References

- 1 A. S. Neish, *Gastroenterology*, 2009, 136, 65–80.
- 2 F. Bäckhed, R. E. Ley, J. L. Sonnenburg, D. A. Peterson and J. I. Gordon, *Science (80-.)*, 2005, 307, 1915–1920.
- 3 B. Yang, L. Cao, B. Liu, C. D. McCaig and J. Pu, *PLoS One*, 2013, 8, e60861.
- 4 K. B. Martinez, V. Leone and E. B. Chang, *J. Biol. Chem.*, 2017, 292, 8553–8559.
- 5 A. McIntyre, P. R. Gibson and G. P. Young, *Gut*, 1993, 34, 386–391.
- 6 D. L. Lazarova, C. Chiaro and M. Bordonaro, *BMC Res. Notes*, 2014, 7, 226.
- 7 R. Han, Q. Sun, J. Wu, P. Zheng and G. Zhao, *Cell. Physiol. Biochem.*, 2016, 39, 1919–1929.
- 8 L. Zuo, M. Lu, Q. Zhou, W. Wei and Y. Wang, *Food Chem. Toxicol.*, 2013, 62, 892–900.
- 9 A. Koh, F. De Vadder, P. Kovatcheva-Datchary and F. Bäckhed, *Cell*, 2016, 165, 1332–1345.
- 10 L. E. M. Willemsen, M. A. Koetsier, S. J. H. Van Deventer and E. A. F. Van Tol, *Gut*, 2003, 52, 1442–1447.
- 11 H. J. Flint, K. P. Scott, P. Louis and S. H. Duncan, *Nat. Rev. Gastroenterol. Hepatol.*, 2012, 9, 577–589.
- 12 L. Klepinina, A. Klepinin, L. Truu, V. Chekulayev, H. Vija, K. Kuus, I. Teino, M. Pook, T. Maimets and T. Kaambre, *PLoS One*, 2021, 16, e0245348.
- 13 E. Gaudier, M. Rival, M. P. Buisine, I. Robineau and C. Hoebler, *Physiol. Res.*, 2009, 58, 111–119.
- 14 Y. P. Silva, A. Bernardi and R. L. Frozza, *Front. Endocrinol. (Lausanne)*, 2020, 11, 25.
- 15 G. M. H. Birchenough, M. E. V. Johansson, J. K. Gustafsson, J. H. Bergström and G. C. Hansson, *Mucosal Immunol.*, 2015, 8, 712–719.
- 16 J. P. M. Van Putten and K. Strijbis, *J. Innate Immun.*, 2017, 9, 281–299.
- 17 P. Paone and P. D. Cani, *Gut*, 2020, 69, 2232–2243.
- 18 K. S. B. Bergstrom and L. Xia, *Glycobiology*, 2013, 23, 1026–1037.
- 19 J. Amano, K. Kobayashi and M. Oshima, *Arch. Biochem. Biophys.*, 2001, 395, 191–198.
- 20 A. Burgess, S. R. Wainwright, L. S. Shihabuddin, U. Rutishauser, T. Seki and I. Aubert, *Dev. Neurobiol.*, 2008, 68, 1580–1590.
- 21 E. Gaudier, L. Forestier, V. Gouyer, G. Huet, R. Julien and C. Hoebler, *Biochem. Biophys. Res. Commun.*, 2004, 325, 1044–1051.
- 22 S. Shah, P. Lance, T. J. Smith, C. S. Berenson, S. A. Cohen, P. J. Horvath, J. T. Y. Lau and H. Baumann, *J. Biol. Chem.*, 1992, 267, 10652–10658.
- 23 T. H. Jung, J. H. Park, W. M. Jeon and K. S. Han, *Nutr. Res. Pract.*, 2015, 9, 343–349.
- 24 S. Cornick, A. Tawiah and K. Chadee, *Tissue Barriers*, 2015, 3.
- 25 K. Steliou, M. S. Boosalis, S. P. Perrine, J. Sangerman and D. V. Faller, *Biores. Open Access*, 2012, 1, 192–198.
- 26 M. Y. Lin, M. R. De Zoete, J. P. M. Van Putten and K. Strijbis, *Front. Immunol.*, 2015, 6, 3.
- 27 T. Buhrke, I. Lengler and A. Lampen, *Dev. Growth Differ.*, 2011, 53, 411–426.
- 28 R. Stierum, M. Gaspari, Y. Dommels, T. Ouatas, H. Pluk, S. Jespersen, J. Vogels, K. Verhoeckx, J. Groten and B. Van Ommen, *Biochim. Biophys. Acta - Proteins Proteomics*, 2003, 1650, 73–91.

- 29 A. V. Pshezhetsky, M. Fedjaev, L. Ashmarina, A. Mazur, L. Budman, D. Sinnett, D. Labuda, J.-F. F. Beaulieu, D. Ménard, I. Nifant'ev and É. Levy, *Proteomics*, 2007, 7, 2201–2215.
- 30 M. Tadjali, J. B. Seidelin, J. Olsen and J. T. Troelsen, *Biochim. Biophys. Acta - Mol. Cell Res.*, 2002, 1589, 160–167.
- 31 N. Turck, S. Richert, P. Gendry, J. Stutzmann, M. Keding, E. Leize, P. Simon-Assmann, A. Van Dorsselaer and J. F. Launay, *Proteomics*, 2004, 4, 93–105.
- 32 E. Tremblay, J. Auclair, E. Delvin, E. Levy, D. Ménard, A. V. Pshezhetsky, N. Rivard, E. G. Seidman, D. Sinnett, P. H. Vachon and J. F. Beaulieu, *J. Cell. Biochem.*, 2006, 99, 1175–1186.
- 33 G. Xu, E. Goonatilleke, S. Wongkham and C. B. Lebrilla, *Anal. Chem.*, 2020, 92, 3758–3768.
- 34 K. Madunić, T. Zhang, O. A. Mayboroda, S. Holst, K. Stavenhagen, C. Jin, N. G. Karlsson, G. S. M. Lageveen-Kammeijer and M. Wuhrer, *Cell. Mol. Life Sci.*, 2020, 1, 3.
- 35 S. Holst, A. J. M. Deuss, G. W. Van Pelt, S. J. Van Vliet, J. J. Garcia-Vallejo, C. A. M. Koeleman, A. M. Deelder, W. E. Mesker, R. A. Tollenaar, Y. Rombouts and M. Wuhrer, *Mol. Cell. Proteomics*, 2016, 15, 124–140.
- 36 R. B. McComb, G. N. J. Bowers and S. Posen, *Alkaline Phosphatase - Robert B. McComb, George N. Bowers, Jr., Solomon Posen - Google Books*, 1979.
- 37 U. Möglinger, S. Grunewald, R. Hennig, C. W. Kuo, F. Schirmeister, H. Voth, E. Rapp, K. H. Khoo, P. H. Seeberger, J. C. Simon and D. Kolarich, *Front. Oncol.*, 2018, 8, 70.
- 38 A. V. SAVAGE, J. J. DONOHUE, C. A. M. KOELEMAN and D. H. van den EIJDEN, *Eur. J. Biochem.*, 1990, 193, 837–843.
- 39 A. V. Savage, S. M. T. D'Arcy and C. M. Donoghue, *Biochem. J.*, 1991, 279, 95–103.
- 40 A. V. SAVAGE, C. M. DONOGHUE, S. M. D'ARCY, C. A. M. KOELEMAN and D. H. van den EIJDEN, *Eur. J. Biochem.*, 1990, 192, 427–432.
- 41 P. A. Prieto, R. D. Larsen, M. Cho, H. N. Rivera, A. Shilatfard, J. B. Lowe, R. D. Cummings and D. F. Smith, *J. Biol. Chem.*, 1997, 272, 2089–2097.
- 42 P. Link-Lenczowski, M. Jastrzębska, K. Chwalenia, M. Pierzchalska, A. Leja-Szpak, J. Bonior, P. Pierzchalski and J. Jaworek, *Biochim. Biophys. Acta - Mol. Cell Res.*, 2019, 1866, 118555.
- 43 F. Qi, T. Isaji, C. Duan, J. Yang, Y. Wang, T. Fukuda and J. Gu, *FASEB J.*, 2020, 34, 881–897.
- 44 R. Pinto, R. Barros, I. Pereira-Castro, P. Mesquita, L. T. Da Costa, E. P. Bennett, R. Almeida and L. David, *Lab. Investig.*, 2015, 95, 718–727.
- 45 N. Carmona-Vicente, D. J. Allen, J. Rodríguez-Díaz, M. Iturriza-Gómara and J. Buesa, *Virol. J.*, DOI:10.1186/s12985-016-0538-y.
- 46 K. Murakami, C. Kurihara, T. Oka, T. Shimoike, Y. Fujii, R. Takai-Todaka, Y. Bin Park, T. Wakita, T. Matsuda, R. Hokari, S. Miura and K. Katayama, *PLoS One*, 2013, 8, e66534.
- 47 J. Amano and M. Oshima, *J. Biol. Chem.*, 1999, 274, 21209–21216.
- 48 E. Walter and T. Kissel, *Eur. J. Pharm. Sci.*, 1995, 3, 215–230.
- 49 N. Malagolini, F. Dall'Olio and F. Serafini-Cessi, *Biochem. Biophys. Res. Commun.*, 1991, 180, 681–686.
- 50 H. R. Wang, C. Y. Hsieh, Y. C. Twu and L. C. Yu, *Glycobiology*, 2008, 18, 104–113.
- 51 X. Sun and M. J. Zhu, *Mol. Nutr. Food Res.*, 2018, 62, e1700932.
- 52 C. Sato and K. Kitajima, *J. Biochem.*, 2013, 154, 115–136.
- 53 X. Zhou, G. Yang and F. Guan, *Cells*, 2020, 9, 273.

- 54 M. Teinténier-Lelièvre, S. Julien, S. Juliant, Y. Guerardel, M. Duonor-Cérutti, P. Delannoy and A. Harduin-Lepers, *Biochem. J.*, 2005, 392, 665–674.
- 55 A. Fernández-Briera, I. García-Parceiro, E. Cuevas and E. Gil-Martín, *Oncology*, 2010, 78, 196–204.
- 56 J. Li, H. C. Hsu, J. D. Mountz and J. G. Allen, *Cell Chem. Biol.*, 2018, 25, 499–512.
- 57 T. Zhang, K. Madunić, S. Holst, J. Zhang, C. Jin, P. Ten Dijke, N. G. Karlsson, K. Stavenhagen and M. Wuhler, *Mol. Omi.*, 2020, 16, 355–363.

Supplementary Figures

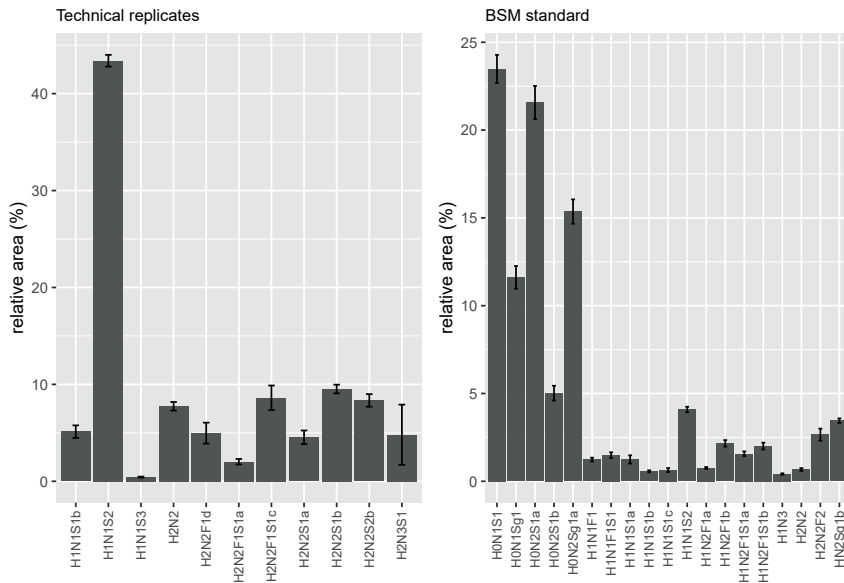


Figure S1 | Robustness of the glycomic workflow a) Variability of the sample preparation workflow for three technical replicates from the same cell lysate (5 days – spontaneous differentiation) b) Variability of 5 measurements of bovine submaxillary mucin released *O*-glycans across 5 days.

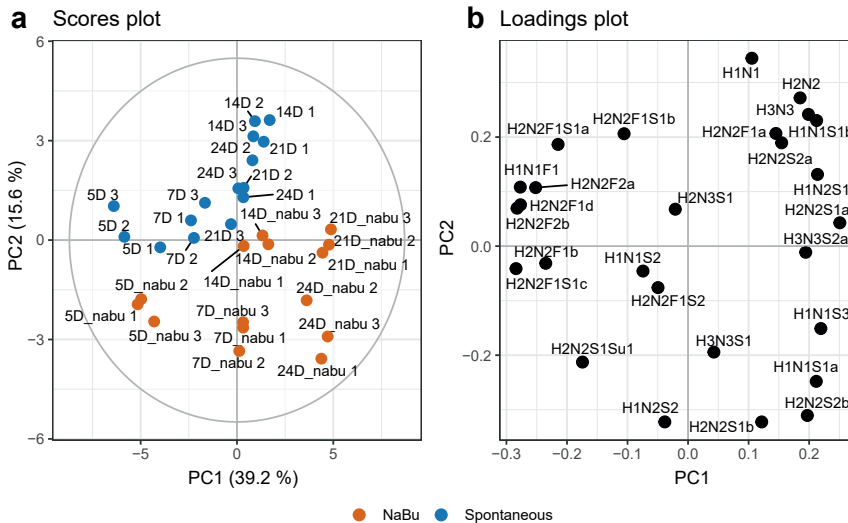


Figure S2 | PCA model based on relative abundances (%) of released *O*-glycans. a) PCA scores plot with 3 biological replicates (labeled 1, 2 or 3) per condition (nabu/none) and per timepoint (5D, 7D, 14D, 21D, and 24D). The top two principal components explain 55.1 % of the variation within the data. b) The PCA loadings plot displays the variables that drive the separation in the PCA model

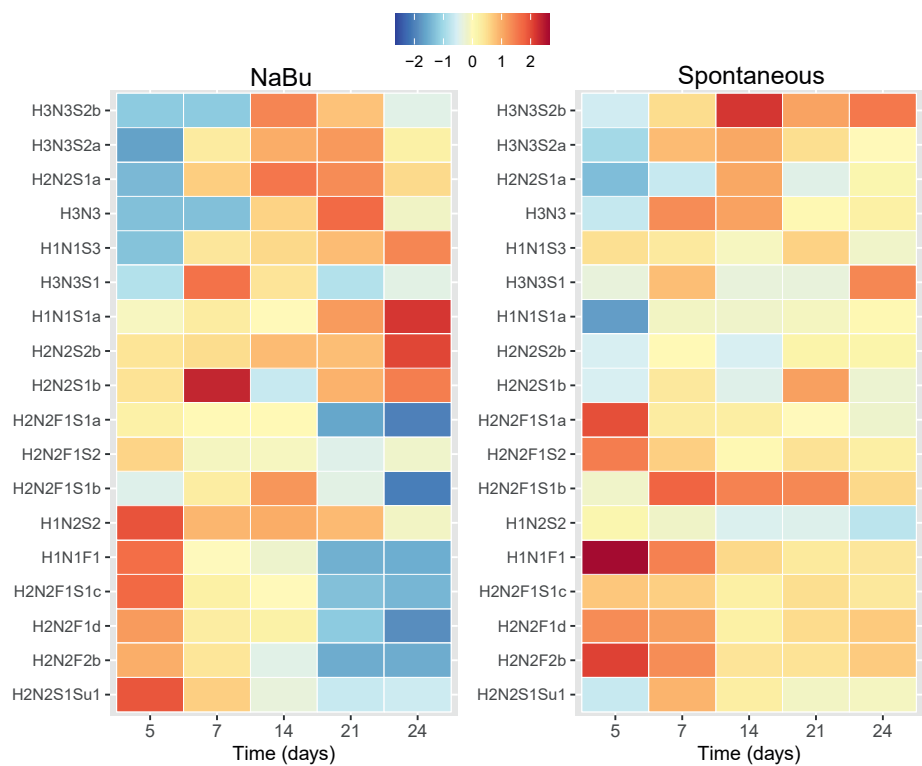


Figure S3 | Relative abundance of O-glycans that show a significant difference with differentiation. Geometrical tile of the scaled relative abundances (%) of individual glycans (y-axis) across different timepoints (x-axis) selected from the analysis of variance (ANOVA) with NaBu stimulation and without (spontaneous differentiation).

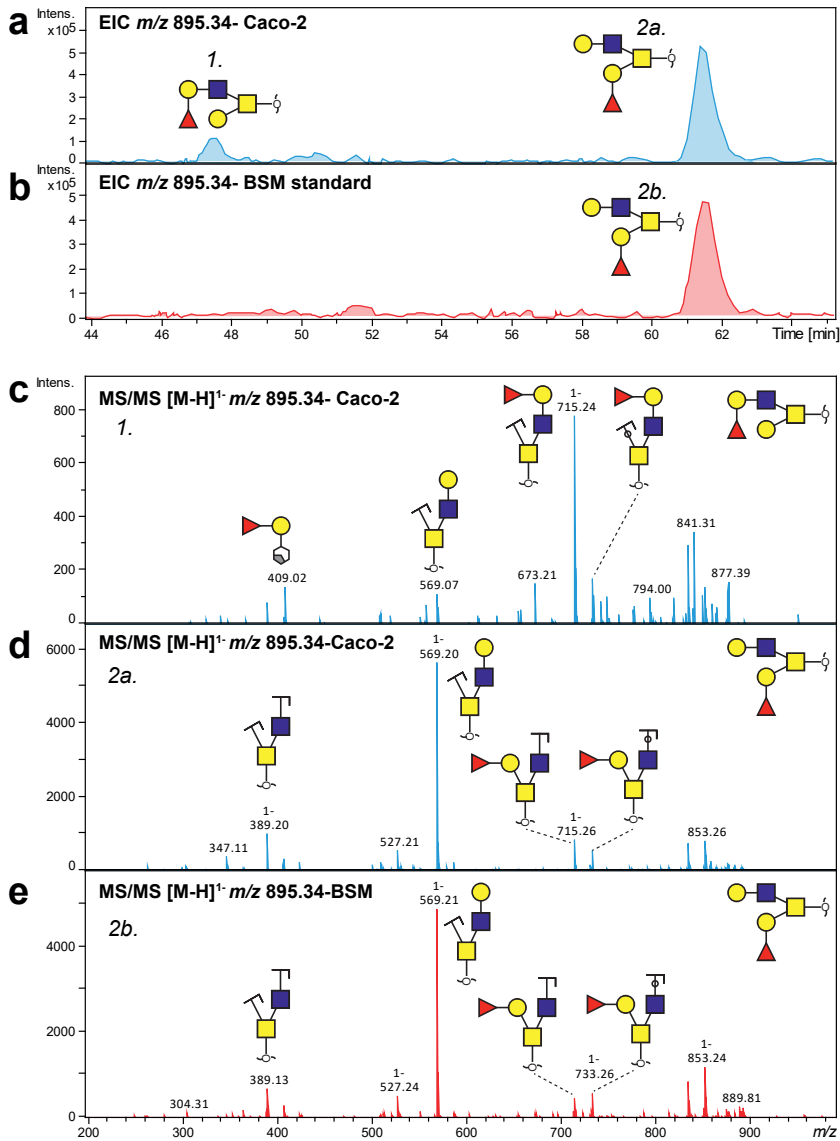


Figure S4 | Structural determination of the O-glycan isomers with composition H2N2F1. Extracted ion chromatograms (EICs) of m/z 895.34 corresponding to the O-glycan mass of H2N2F1 in the (a) Caco-2 cell line sample and (b) bovine submaxillary mucin (BSM) standard. The Caco-2 cell line sample reveals two isomeric species while the BSM standard contains a single O-glycan. The structures of the isomers eluting at 47.5 and 61.8 min could be assigned based upon the obtained MS/MS spectra. (c) The presence of the characteristic cross-ring fragment indicates a type 2 blood group H antigen on the 6-arm, which is further supported by the abundant Z ion at m/z 715.2 for the first isomer in the Caco-2 cell line. (d and e) The presence of the Z fragment ions at m/z 389.2 and 569.2 indicate an arm occupancy on the 6-arm, carrying a terminal α 1,2-linked fucose to the galactose on 3-arm. This isomer has been identified previously as the major isomer in BSM by NMR.

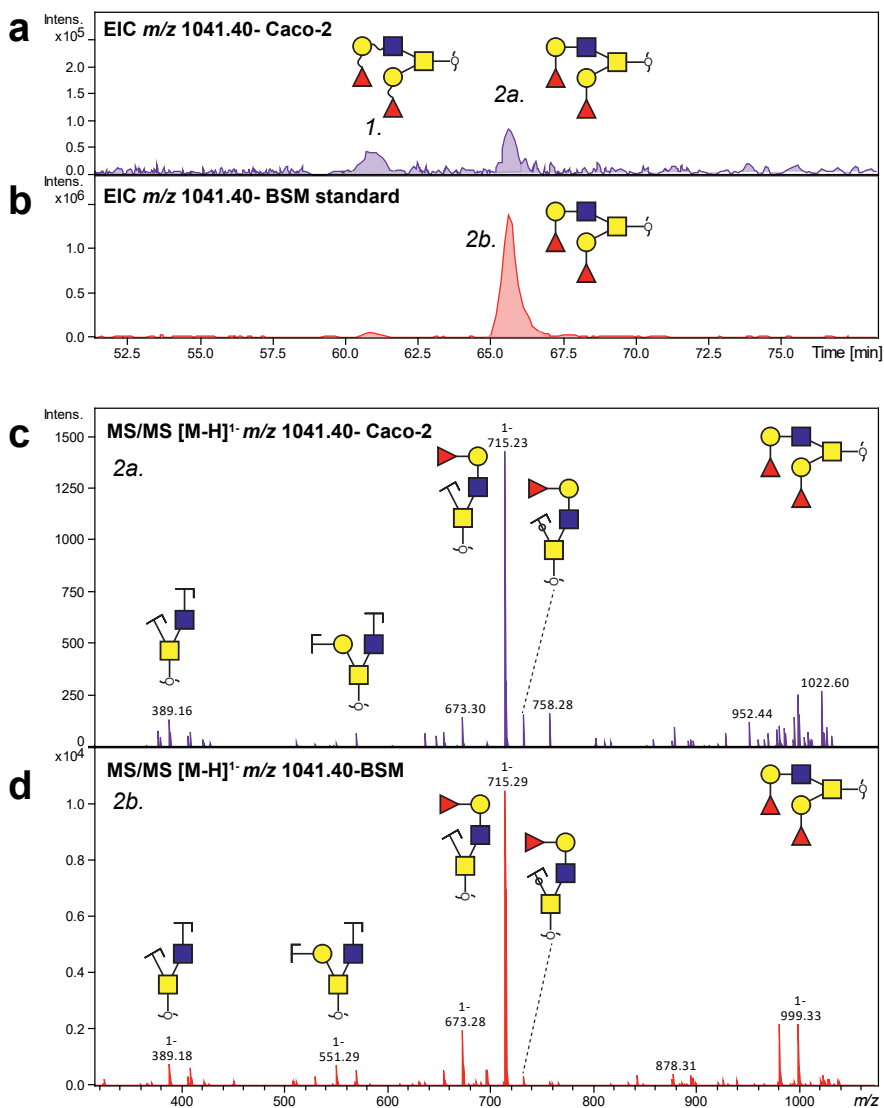


Figure S5 | Structural determination of the *O*-glycan isomers with composition H2N2F2. Extracted ion chromatograms (EICs) of m/z 1041.40 corresponding to the *O*-glycan mass H2N2F2 in the (a) Caco-2 cell line sample and (b) bovine submaxillary mucin (BSM) standard. The Caco-2 cell line sample reveals two isomeric species while the BSM standard contains a single *O*-glycan. The structure of the isomer eluting at 65.5 min could be assigned based upon the obtained MS/MS spectra. (c and d) The presence of the characteristic Y- and Z-fragment ion pairs at m/z 733.3 and 715.3, respectively indicate an arm occupancy on the 6-arm, carrying a terminal α 1,2-linked fucose on both arms. This isomer has been identified previously as the major isomer in BSM by NMR. Due to poor quality of the MS/MS spectra, as well as no corresponding isomer in the BSM standard, the linkages in isomer 1 could not be determined with high confidence.

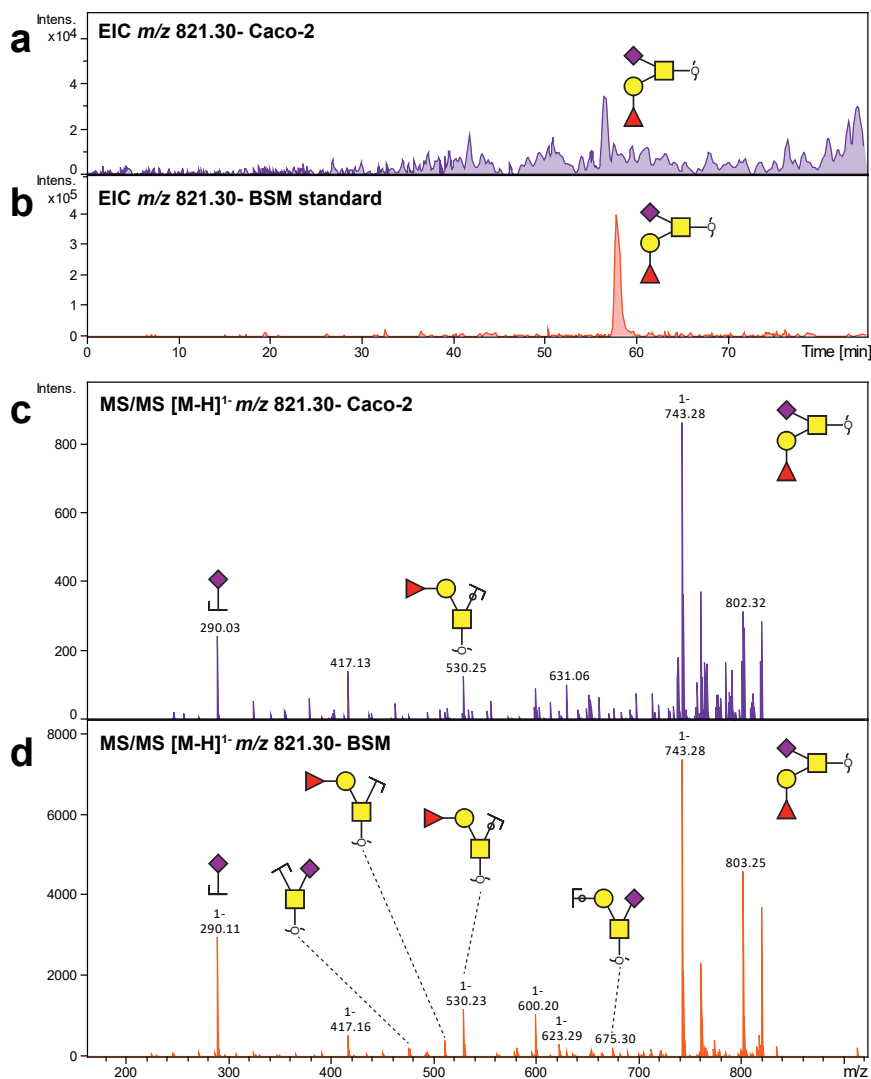


Figure S6 | Structural determination of the O-glycan with composition H1N1F1S1. Extracted ion chromatograms (EICs) of m/z 821.30 corresponding to the O-glycan mass H1N1F1S1 in the (a) Caco-2 cell line sample and (b) bovine submaxillary mucin (BSM) standard. The structure of the isomer eluting at 59.2 min could be assigned based upon the obtained MS/MS spectra. (c and d) The presence of the characteristic Y and Z fragment ion pairs at m/z 512.1 and 530.2, respectively as well as Z ion at m/z 495.25 indicate the sialic acid linkage to the innermost N-acetylglucosamine, and a type 3 blood group antigen H on the 3-arm.

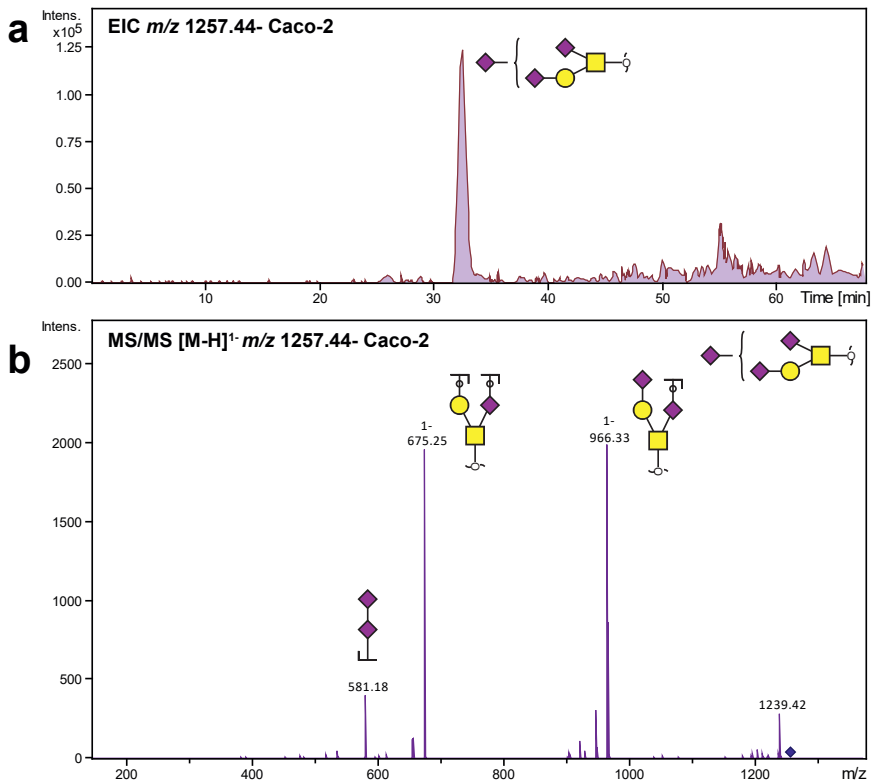
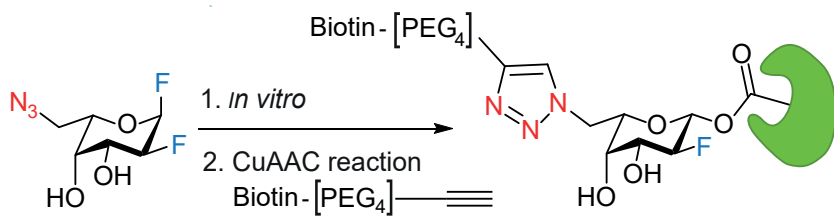


Figure S7 | Structural determination of the O-glycan with composition H1N1S3. Extracted ion chromatogram (EICs) of m/z 1257.44 corresponding to the O-glycan mass H1N1S3 in the **(a)** Caco-2 cell line sample. The structure eluting at 33.2 min was assigned based upon the obtained MS/MS spectra. **(b)** The presence of the characteristic B-ion at m/z 581.2 indicate the presence of two sialic acids linked together. It could not be determined whether the two sialic acids are present on the 6-arm or the 3-arm of the glycan.



- ✓ Readily accessible synthesis
- ✓ **Azide** allows functionalization with different reporter groups

- ✓ Detect **fucosidases** with different substrate specificity

Chapter 4

Development of a 1,2-difluorofucoside activity-based probe for profiling GH29 fucosidases

Yvette. M.C.A. Luijkx, Seino A.K. Jongkees, Karin Strijbis, Tom Wennekes

Adapted from: Organic & Biomolecular Chemistry, 2021, DOI:10.1039/D1OB00054C

Abstract

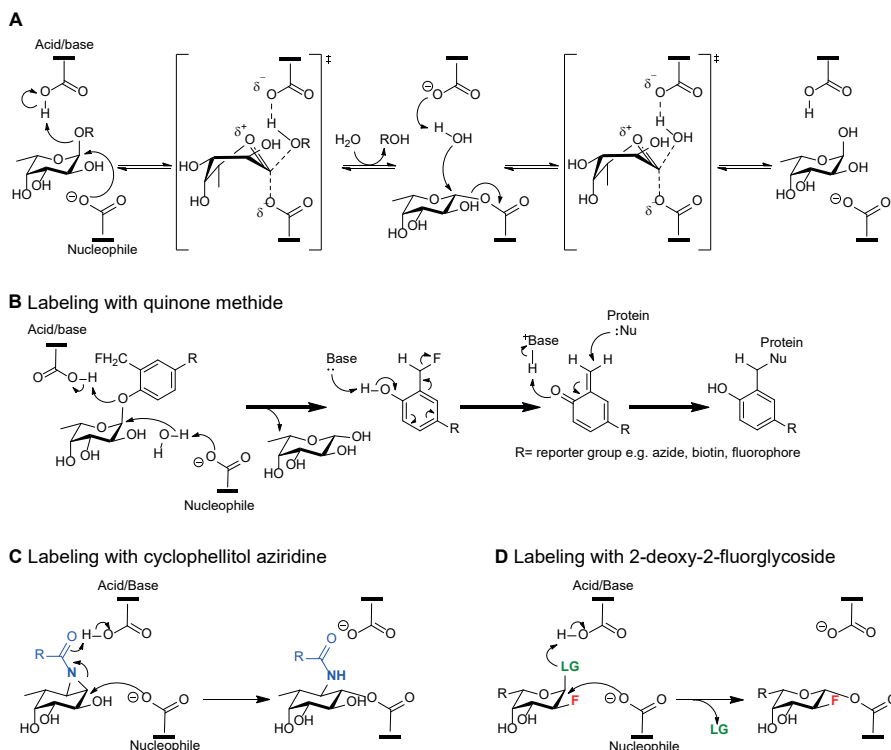
GH29 α -L-fucosidases catalyze hydrolysis of terminal α -L-fucosyl linkages with varying specificity and are expressed by prominent members of the human gut microbiota. Both homeostasis and dysbiosis at the human intestinal microbiota interface have been correlated with altered fucosidase activity. Herein we describe the development of a 2-deoxy-2-fluoro fucosyl fluoride derivative with an azide mini-tag as an activity-based probe (ABP) for selective *in vitro* labelling of GH29 α -L-fucosidases. Only catalytically active fucosidases are inactivated by this ABP, allowing their functionalization with a biotin reporter group via the CuAAC reaction and subsequent in-gel detection at nanogram levels. The ABP we present here is shown to be active against a GH29 α -L-fucosidases from *Bacteroides fragilis* and capable of labeling two other GH29 α -L-fucosidases with different linkage specificity, illustrating its broader utility. This novel ABP is a valuable addition to the toolbox of fucosidase probes by allowing identification and functional studies of the wide variety of GH29 fucosidases, including those in the gut microbiota.

Introduction

Retaining α -L-fucosidases (GH29 fucosidases) are exo-glycosidases that catalyze the hydrolysis of terminal α -L-fucopyranosides from cellular glycans.¹ This glycosidase family is widespread in bacteria, fungi and mammals, and implicated in several critical biological processes such as immune response, signal transduction and adhesion of pathogens.^{1–3} GH29 fucosidases have a diverse substrate specificity. They commonly hydrolyze α -(1,2) linkages from fucose to galactose or α -(1,3), α -(1,4), and α -(1,6) linkages to *N*-acetylglucosamine. Based on their substrate specificity they can be divided in two main subfamilies (A and B).⁴ Subfamily A fucosidases have relatively relaxed substrate specificity. For example the human GH29 fucosidase is assigned to subfamily A as it is active on α -1,2/3/4/6-linked fucosyl substrates.⁵ Similar to this the GH29 fucosidase Bt2970 from the intestinal bacterium *Bacteroides thetaiotaomicron* shows a broad substrate specificity.⁶ In contrast, subfamily B is more specific towards α -(1,3) or α -(1,4) linkages. Several other GH29 fucosidases, among which BT2192 and BT3798 (from *B. thetaiotaomicron*) show a more stringent specificity, placing them in subfamily B.^{4,6,7} We are interested in elucidating the role of this diverse set of bacterial fucosidases specifically in the human gut where they have been implicated in host-microbe interplay, microbiota cross feeding and colonization resistance.^{8–10} For example, it has been established that infection efficiency of the common foodborne pathogens, *Salmonella typhimurium* and *Campylobacter jejuni*, relies on fucosidases secreted by other residing bacteria.^{2,11,12} To study the exact function of fucosidases produced by pathogens and commensals we first have to identify them and determine their spatial and temporal distribution.¹³ Investigations into the functions of different bacterial fucosidases will benefit from the availability of chemical probes that enable the screening, isolation and study of catalytically-active fucosidases.

GH29 fucosidases hydrolyze their substrate with overall retention of configuration at the anomeric center through a classic Koshland double displacement mechanism (Scheme 1A).¹⁴ The reaction mechanism involves two carboxylates from aspartic acids in the active site, one acting as the catalytic nucleophile and the other acting as the catalytic acid-base. The mechanism proceeds through a covalent fucosyl-enzyme intermediate accessed via a pair of oxocarbenium ion-like transition states. Quinone methide based probes for glycosidases have been reported to generate a reactive electrophile upon hydrolysis of their glycosidic bond, of which one example covalently labeled a human fucosidase (Scheme 1B).^{15–18} Although these probes have shown their merit in several studies, their major limitation is that the affinity for the enzyme is lost by cleavage of the glycosidic bond resulting in possible non-specific labeling of non-associated proteins. 2-Deoxy-2-fluoro glycosides and cyclophellitol epoxides or aziridines are the two main classes of covalent inhibitors that, opposed to the quinone methide based probes, can inactivate and label a glycosidase by trapping the covalent enzyme intermediate.^{19,20} By altering the stereochemistry, substituents and warhead of the cyclophellitol scaffold,

Overkleeft and co-workers have developed ABPs for a wide range of different retaining glycosidases.^{21–24} A study by Davies and co-workers provided the first structural basis for the inactivation of a β -glycosidase by cyclophellitol, employing an electrophilic trap mechanism to label the target enzyme (Scheme 1C).²⁵



Scheme 1 | (A) Classic Koshland double displacement mechanism for GH29 fucosidases. Labeling mechanism of a fucosidase for a quinone methide based probe (B), a cyclophellitol aziridine based probe (C) and a 2-deoxy-2-fluoro fucoside-based probe (D).

2-Deoxy-2-fluoro glycosides were originally introduced by Withers and co-workers.^{26,27} The substitution of a highly electronegative fluorine atom for the C2 hydroxyl group destabilizes the oxocarbenium ion-like states, resulting in a decreased rate for both the formation of the glycosyl-enzyme intermediate and its subsequent hydrolysis. By employing a suitably reactive leaving group at the anomeric site the formation of the covalent intermediate will still proceed. Essentially, these act as a very slow substrate and result in accumulation of the glycosyl-enzyme intermediate (Scheme 1D). Activated 2-deoxy-2-fluoro glycosides have been used to identify and label catalytic nucleophiles of a variety of retaining glycosidases and this class of inhibitors has been extensively reviewed over the past decades.^{20,21,28–30} The inhibition of specific glycosidases by 2-deoxy-2-fluoro glycoside-based inhibitors has been demonstrated in live animals, and

showed promise in future imaging applications by applying radiolabelled inactivators.^{31,32} Furthermore, activated 2-deoxy-2-fluoro glycosides were the basis for activity-based protein profiling (ABPP) probes to target β -D-glucosaminidases, β -galactosidases, and β -glycanases, when conjugated to a reporter group such as biotin or a fluorophore.^{20,33–36} The application in this technique allowed for the first successful selective labelling of β -galactosidases in a complex proteome.³⁴

In 2003 Withers *et al.* set a precedent for mechanism-based inhibitors of retaining fucosidase as they reported 2-deoxy-2-fluoro-fucosyl fluoride as a selective retaining fucosidase inhibitor and identified the catalytic nucleophile of the GH29 fucosidase of the thermophilic bacterium *Thermogata maritima*.³⁷ This was followed in 2015 by Overkleeft *et al.* who developed fucopyranose-configured cyclophellitol aziridine ABPs for selective labelling of GH29 fucosidases in bacteria, mice and man.²² Although the reported cyclophellitol based ABPs have shown to be more potent inhibitors compared to their 2-deoxy-2-fluoro-glycoside analogues, their multi-step synthesis remains a major hurdle.

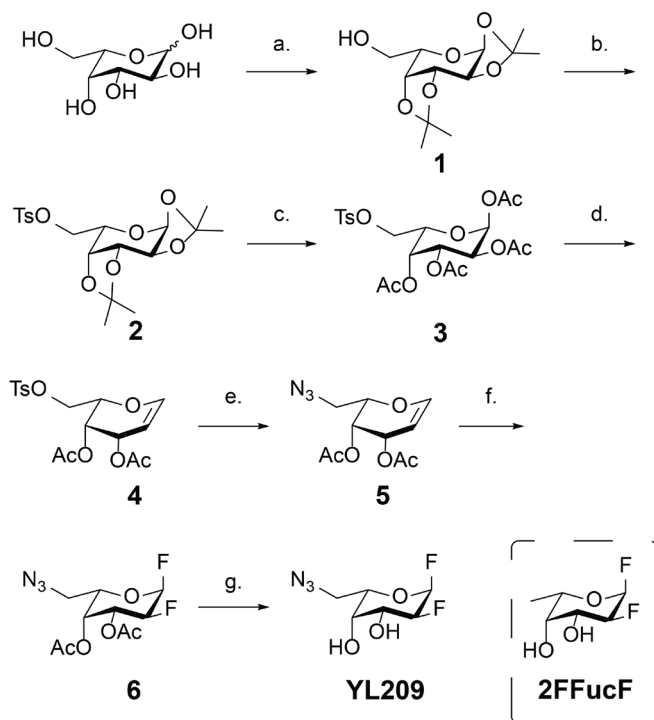
One of the long-term aims within our research program is to identify the role of bacterial fucosidases in both homeostasis and dysbiosis in the human gut. We have previously reported that *Campylobacter jejuni fuc+* strains are dependent on exogenous microbial fucosidases for increased growth and invasion.¹² This highlights the importance of fucosidases in the development of infections and shows the merit for the development of different ABPs to further study these processes and identify the responsible enzymes and organisms. In this study our aim was therefore to investigate if we could transform the synthetically more readily accessible 2-deoxy-2-fluor-fucosides into an ABP that can detect microbial α -L-fucosidases in a time- and concentration-dependent manner.

We describe here the successful synthesis of a 2-deoxy-2-fluoro-fucosyl fluoride-based ABP (Scheme 2, **YL209**) compatible with click chemistry via an azide mini-tag and demonstrate the versatility of this probe by applying it to GH29 fucosidases with varying linkage specificity.

Results and discussion

Our first aim was to convert the 2-deoxy-2-fluoro-fucosyl fluoride inhibitor into an ABP by developing a synthetic route that allows for the introduction of a functional group that can be used to form a covalent linkage to a reporter group such as a biotin or fluorophore. There is a precedent for fucosides with substituents at the 6-position to serve as a substrate for different fucosidases³⁸, and various other glycosidases also seem to tolerate different substitutions at the 6-position.^{39–41} Therefore we decided to introduce an azide mini-tag on the C6 position of the fucoside. This position additionally retains free rotation around the C6 methylene group that may permit the azide to adopt

a favorable position in the fucosidase active site, which is necessary in downstream clicking to the reporter group.



Scheme 2 | Synthesis of **YL209**, in which the 2FFucF structure is also shown. *Reagents and conditions:* (a) ZnCl_2 , cat. H_2SO_4 , acetone, 85%; (b) TsCl , DMAP, pyridine, 0°C to rt, 86%; (c) *i.* TFA/ H_2O (1:1); *ii.* Ac_2O , pyridine, 92%; (d) *i.* HBr (37% in AcOH), DCM 0°C to rt; *ii.* sat. aq. NaH_2PO_4 , Zn , Acetone, 77%; (e) NaN_3 , DMF, 90°C , 95%; (f) XeF_2 , $\text{BF}_3\cdot\text{OEt}_2$, benzene, diethyl ether, -10°C to rt, 57%; (g) NaOMe , MeOH , 80%.

We aimed to synthesize the ABP via the fluorination of a glycal intermediate (Scheme 2) starting from L-galactose. First, all secondary hydroxy groups were protected as di-isopropylidene acetals to afford **1** in 85% yield. The free C6 hydroxyl in intermediate **1** was converted into C6 tosylate **2**. Next, the di-isopropylidene acetals were hydrolyzed and the liberated hydroxyls protected as acetyl esters to provide compound **3** in 79% yield over the three steps. Our initial approach to obtain the 6-azido-galactal **5** was to perform an azide substitution on the C6 tosylate of **3** followed by anomeric bromination and zinc mediated reductive elimination to the 6-azido glycal. This however did not lead to the desired product, but instead resulted in substitution of the C6 azide by a bromide. We therefore decided to perform the bromination and elimination prior to introducing the azide. Bromination of **3** with HBr gave the anomeric bromide intermediate that was immediately treated with zinc dust following a classical Fisher-Zach reductive elimination to give intermediate glycal **4** in 77% yield. Substitution of the tosylate of **4** with sodium azide at elevated temperatures gave compound **5** in an excellent 95% yield.

Treatment of the protected L-galactal derivative **5** with XeF_2 in the presence of a lewis acid resulted in difluorination of the alkene from the top face to provide an equatorial C2 fluorine and as the main product the favorable α anomeric fluorine. Deacetylation of **6** yielded the desired probe **YL209**. To investigate the effect of the azide at C6 on inactivation of fucosidases we also synthesized the known fucosidase inhibitor 2FFucF, with a methyl group as C6, according to a literature procedure.^{37,42}

The leaving group at C1 of the fluorinated glycosides has previously been shown to play an important role in the effectiveness of these probes as ABPs. Overkleeft *et al.* reported that the installation of an anomeric *N*-phenyl trifluoroacetimidate on 2-fluoro- β -glycopyranose-based ABPs drastically improved their potency.⁴³ We therefore also attempted to synthesize the *N*-phenyl trifluoroacetimidate on our 2-fluoro-6-azido α -L-fucosyl based ABP. Although we could obtain the penultimate C2, C3 levulinoyl ester protected intermediate, in our hands we were unable to obtain the final molecule due to decomposition during deprotection (Figure S1). This result can perhaps be explained by the relatively high instability/reactivity of galactosides and fucopyranosides due to electron withdrawing properties of their axial 4-OH.⁴⁴

With the putative mechanism-based inhibitor **YL209** in hand, we sought to evaluate its ability to inactivate a microbial fucosidase. We selected a fucosidase expressed by the well-known gut commensal, *Bacteroides fragilis*, as the initial target enzyme. The selectivity of BfFucH has already been reported and shows a broad substrate recognition. Incubation of BfFucH with various concentrations of **YL209** revealed a time and concentration-dependent inactivation of the enzyme that could be fit to a single exponential decay equation (Figure 1A). However, incubation of BfFucH with various concentrations **YL209** never resulted in complete inactivation. When incubated with 5, 10, 20, 40, and 80 mM **YL209** respectively 43%, 48%, 60%, 70%, and 76% inactivation could be achieved. This prompted us to investigate whether **YL209** itself was stable under the conditions used. TLC analysis of the probe incubated in PBS over time however showed no degradation of the probe (Supporting Figure 2). We next investigated whether the **YL209** covalent intermediate with the enzyme was hydrolyzing at an appreciable rate, thereby reactivating the enzyme. This would result in activity reaching an equilibrium based on competing rates of inactivation by **YL209** and enzyme reactivation, and thus explaining the incomplete inactivation. BfFucH was incubated with excess **YL209** and initial hydrolysis rates were determined using the 4-nitrophenyl- α -L-fucopyranoside after excess inhibitor was removed (Figure 1B). Time-dependent reactivation of a glycosidase that had been inactivated with a similar difluoro glycoside based probe was described before.⁴⁵ Similar to what was described by Kim and co-workers we observed an increase of the initial hydrolysis rate over time, according to a hyperbolic function, suggesting turnover of the enzyme-intermediate playing a role in the earlier observed equilibrium.

We next investigated the inactivation kinetics at several concentrations of **YL209** and the kinetic parameters (k_{inact} and K_i) were determined by plotting k_{obs} versus $[I]$ (Figure 1C). Inactivation of BfFucH by **YL209** exhibited a good fit to all data by non-linear regression using the equation $K_{\text{obs}} = k_{\text{inact}} / (1 + K_i/[I])$, from which a K_i value of 9.6 ± 2.7 mM and a k_{inact} value of $0.029 \pm 0.002 \text{ min}^{-1}$ could be deduced. We next determined the K_i and k_{inact} for the related known inactivator, 2FFucF (Scheme 2), to be respectively 26.4 ± 10.1 mM and $0.009 \pm 0.001 \text{ min}^{-1}$. The inhibition constant K_i for **YL209** shows it to be a more potent inhibitor compared to 2FFucF, as the concentration needed to produce half maximum inhibition is lower. This is supported by the k_{inact} value indicating a higher rate of inactivation for **YL209** compared to 2FFucF. The second-order rate constant for ABP **YL209** ($k_{\text{inact}}/K_i = 3.0 \pm 0.003 \text{ M}^{-1} \text{ min}^{-1}$) that governs inactivation compared well to 2FFucF ($k_{\text{inact}}/K_i = 0.3 \pm 0.010 \text{ M}^{-1} \text{ min}^{-1}$) (Figure 1D). The azide mini-tag at C-6 thus appears to have a beneficial effect on the inactivation process. These experiments show that **YL209** inactivates BfFucH on a time scale appropriate for practical applications and suggests that this compound can be used to detect GH29 fucosidases in more complex samples.

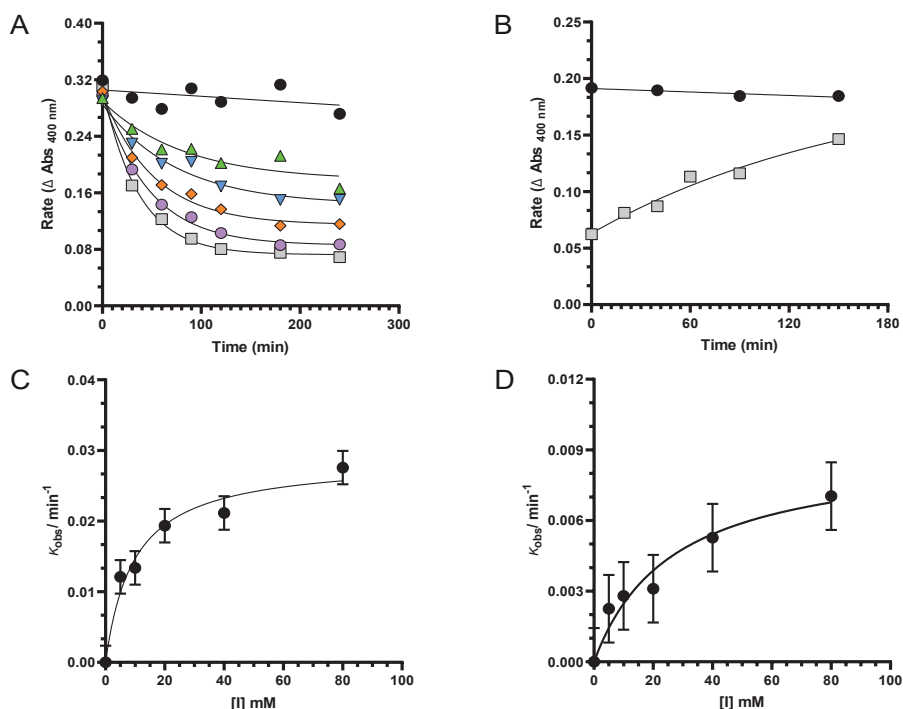


Figure 1 | Time-dependent inactivation of BfFucH. (A) Experimental data represents the BfFucH activity with **YL209**: 0 (●), 5 (▲), 10 (▼), 20 (◆), 40 (●), 80 (■) mM. Curves are the nonlinear fits of data to a single-exponential decay equation. (B) reactivation kinetics for covalently inhibited BfFucH, control samples (absence of inhibitor) (●), pre-incubated with 50 mM **YL209** (■). (C) Plot of the inactivation rate constant (k_{obs}) as function of **YL209** for BfFucH. (D) Plot of the inactivation rate constant (k_{obs}) as function of 2FFucF for BfFucH.

To explore the practical possibilities of the probe to detect fucosidases, we incubated the recombinant BfFucH with **YL209** overnight at 37 °C followed by SDS PAGE gel and electrophoretic transfer to a nitrocellulose blot. For the visualization of the ABP-BfFucH signal we conducted the functionalization with an alkyne-biotin on the surface of the nitrocellulose blot. This method has previously been shown to have similar sensitivity compared to regular imaging methods and it eliminates the need for labor intensive separation steps to remove excess of probe before attachment of the visualization handle.⁴⁶ Incubation of **YL209** with BfFucH followed by SDS-PAGE and Western Blot analysis revealed no signal when normal electrophoresis protocols were followed. It has previously been discussed in literature that this result can occur due to hydrolysis of the ester linkage of the glycosyl-enzyme intermediate under the basic conditions (pH 8.8) of the buffered gel.³³ A short electrophoresis protocol (90 V for 20 minutes and 200 V for 8 minutes) to circumvent this was developed by Stubbs *et al.* However, this protocol was not reproducible in our hands and we therefore selected a different gel system with a more favorable running pH. In contrast to the conventional Tris-glycine SDS-PAGE gels, the BIS-Tris gels have a running pH of 7.2 and this near neutral pH diminishes the reactivity of amino and sulfhydryl groups.⁴⁷

A BIS-TRIS MOPS gel was used in conjunction with the Western Blot analysis above and on blot click reaction showed that these adjusted conditions proved successful for labelling of BfFucH (Figure 2A). To confirm that the alkyne-biotin was reacting specifically with the azide mini-tag of the BfFucH inactivated with **YL209** and not aspecifically with the protein itself, various controls were performed (Figure 2A-B). The results from these controls show that the inactivation and subsequent biotin CuAAc are highly specific and neither reagent on its own results in labelling of the enzyme. Incubating BfFucH with **YL209** revealed the lower limit of detection of BfFuc to be approximately 300 ng (Figure 2C). This is in a similar range as was previously reported for another 2-deoxy-2-fluoro-glycoside ABP.³⁴

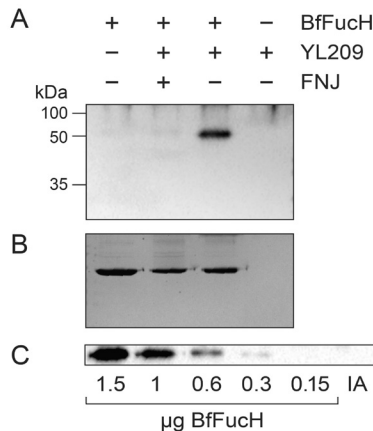


Figure 2 | Visualization of **YL209** labelled BfFucH. After inactivation, the samples were run on BIS-TRIS, blotted on nitrocellulose, labelled by CuAAC on blot with 100 µM alkyne-biotin and analyzed using anti-biotin-HRP. (A) TRIS-BIS-western blot with the appropriate controls. (B) PageBlue stained TRIS-BIS gels, showing equal loading of lanes. (C) Detection limit of **YL209** for specified amount of BfFucH; IA: 1.5 mg BfFucH pre incubated with 100 µM FNJ.

To demonstrate the versatility of our approach we applied **YL209** to two commercially available recombinant GH29 fucosidases. One GH29 fucosidase from a bacterial source, *Thermotoga maritima* (E-FUCTM) with specificity for α -(1,2)-L-fucoside residues and one from a homo sapiens source (E-FUCHS) with a broad specificity for α -(1-2,3,4,6)-L-fucoside residues. In both cases we observed specific labeling of the enzymes when incubated with **YL209** (Figure 3). Noticeable, for E-FUCTM we also see specific labelling of a higher running spot that might correspond to the naturally occurring hexamer conformation of this enzyme.⁴⁸ These results shows that our probe is capable of labeling GH29 fucosidases independent of linkage specificity.

Having established the activity of **YL209**, we looked further into future applications of this probe such as in fucosidase identification in microbiota members. To expand the potential applicability of **YL209** we validated the in-solution labeling of the probe-enzyme complex with a reporter, as opposed to on-blot labeling that we used up to this point. We incubated purified BfFucH with **YL209** overnight at 37°C. The sample was dialyzed to remove excess probe and the protein was denatured under mild reducing conditions at pH 3.5 so that both the azide moiety and the ester linkage of the covalent intermediate remained intact. Treatment of the mixture with alkyne-biotin under Cu^I catalysis followed by Western blot analysis revealed a clear selective labelling of the BfFucH enzyme (Figure 4). In a control we observed that labeling could be lowered by the presence of the active site competitor iminosugar FNJ during incubation with the probe.

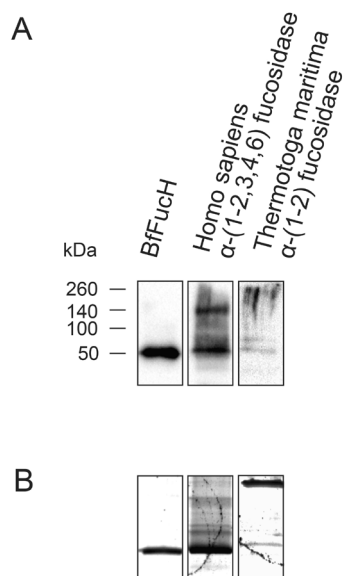


Figure 3 | Labelling of three GH29 fucosidases by using **YL209** in conjunction with in-blot CuAAC. (A) Western blot analysis of samples of recombinant BfFucH (Lane 1), *Homo Sapiens* fucosidase (Lane 2), *Thermotoga maritima* fucosidase (lane 3). (B) TRIS-BIS PageBlue (PB) analysis of the samples shown in (A).

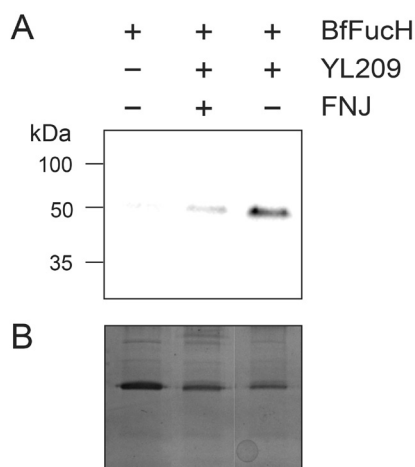


Figure 4 | Western Blot analysis of BfFucH inactivated with **YL209** in conjunction with in-solution CuAAC. (A) TRIS-BIS-western blot with the appropriate controls. (B) TRIS-BIS PageBlue analysis of the samples shown in (A).

Conclusions

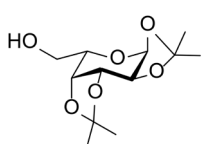
We have successfully synthesized an α -L-fucosidase targeting ABP and developed a strategy to label active retaining fucosidases by using an azide on C6 as a small robust chemical mini-tag. We have demonstrated that the approach can be used to detect several glycosidases with different substrate specificities and with on-blot as well as in-solution labelling. Furthermore, we believe that the readily accessible synthesis described here for **YL209** will pave the way for further investigations into the effect of different C1 or C6 modifications on labeling sensitivity and specificity for α -L-fucosidases. In particular, the leaving group on C1 of the 2-deoxy-2-fluorinated glycosides has previously shown to play an essential role in the effectiveness of these molecules as ABPs, although our attempted installation of an anomeric *N*-phenyltrifluoroacetimidate proved not possible in this case. Together, these properties make this ABP a suitable candidate to further derivatise, identify and study bacterial fucosidases of pathogenic and commensal bacteria in future studies.

Experimental

General methods and materials

All chemicals were obtained from commercial suppliers and used as received unless stated otherwise. Organic solvents for reactions were dried for at least 2 days over molecular sieves (3 or 4 Å). ^1H and ^{13}C NMR spectra were recorded on either an Agilent 400 instrument (400 and 101 MHz) or a Bruker Avance Neo 600 spectrometer (600 and 125 MHz). Chemical shifts are reported in δ values relative to TMS and *J* coupling constants are reported in Hertz (Hz). Silica column chromatography was performed using silica gel SiliaFlash P60 (SiliCycle, Canada, 40–63 μm , 239–400 mesh). TLC analysis was conducted on SiliaPlate TLC Aluminium Backed TLC F254 (SiliCycle) with examination under UV light (254 nm) where applicable, and with 5% sulfuric acid in ethanol or with ceric ammonium molybdate, followed by heating. 2-Deoxy-2-fluoro- α -L-fucosyl fluoride (**2FFucF**) was synthesized according to known methodology.⁴²

1,2,3,4-di-*O*-isopropylidene- α -L-galactopyranose **1**

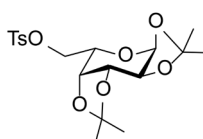


To a solution of ZnCl_2 (800 mg, 5.92 mmol, 2.1 equiv) in acetone (15 mL) was added catalytic H_2SO_4 (27 μL , 0.5 mmol) and L-galactose (500 mg, 2.80 mmol, 1 equiv) at room temperature. The mixture was stirred at room temperature for 5h, and then quenched with sat. aq. NaHCO_3 . The resulting suspension was filtrated over a pad of Celite that was washed twice with acetone. The filtrate was concentrated under reduced pressure and extracted with diethylether. The combined organic layers were dried over sodium sulfate and concentrated under reduced pressure. The residue was purified by silica flash chromatography (7/3 =PE/EtOAc) to afford **1** (620 mg, 2.38 mmol)

as a colorless oil in a 85% yield. R_f = (PE/EtOAc: 8/2 = 0.14); ^1H NMR (400 MHz, CDCl_3) δ 5.51 (d, J = 5.0 Hz, 1H), 4.56 (dd, J = 7.9, 2.4 Hz, 1H), 4.28 (dd, J = 5.0, 2.4 Hz, 1H), 4.22 (dd, J = 8.0, 1.8 Hz, 1H), 3.90 – 3.54 (m, 3H), 1.48 (s, 3H), 1.40 (s, 3H), 1.28 (d, J = 1.6 Hz, 6H). ^{13}C NMR (101 MHz, CDCl_3) δ 109.83, 109.03, 96.65, 71.97, 71.12, 70.93, 68.42, 62.71, 26.37, 26.28, 25.28, 24.65.

Identical with reported data.⁴⁹

1,2,3,4-di-*O*-isopropylidene-6-*O*-(*p*-toluenesulfonyl)- α -L-galactopyranose **2**

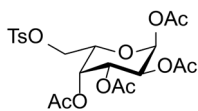


A solution of *p*-toluenesulfonyl chloride (650 mg, 3.41 mmol, 1.8 equiv) in DCM (1 mL) was added to a cooled (-5 °C) solution of **1** (500 mg, 1.92 mmol, 1 equiv) and 4-(dimethylamino)pyridine (40 mg, 0.34 mmol, 0.18 equiv) in anhydrous pyridine (5 mL). The solution was warmed to room temperature and stirred overnight.

The solvent was removed under reduced pressure and the residue dissolved in water and extracted with EtOAc (three times). The organic layer was washed with 0.1 M HCl, saturated bicarbonate, and brine, and dried over sodium sulfate, filtered and evaporated. The residual syrup was purified by silica flash chromatography (8/2 = PE/EtOAc) to afford **2** (685 mg, 1.65 mmol) as a yellow oil in a 86% yield. R_f = (PE/EtOAc: 8/2 = 0.25); ^1H NMR (400 MHz, CDCl_3) δ 7.84 – 7.30 (m, 4H), 5.45 (d, J = 4.9 Hz, 1H), 4.59 (dd, J = 7.9, 2.6 Hz, 1H), 4.29 (dd, J = 5.0, 2.5 Hz, 1H), 4.23 – 4.19 (m, 2H), 4.10 – 4.01 (m, 2H), 2.44 (s, 3H), 1.50 (s, 3H), 1.34 (s, 3H), 1.31 (s, 3H), 1.28 (s, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 145.08, 133.20, 130.08, 128.47, 109.92, 109.29, 96.47, 70.87, 70.75, 70.72, 68.52, 66.21, 26.32, 26.15, 25.26, 24.69, 21.97. HR MS (m/z) calcd for $\text{C}_{19}\text{H}_{26}\text{O}_8\text{S}$ [$\text{M} + \text{Na}$]⁺, 437.1192; found, 437.1243.

Identical with reported data.⁵⁰

1,2,3,4-Tetra-*O*-acetyl-6-*O*-(*p*-toluenesulfonyl)- α -L-galactopyranose **3**

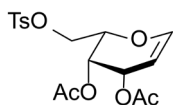


Compound **2** (800 mg, 1.93 mmol, 1 equiv) was dissolved in DCM (2 mL) and added dropwise to a vigorously stirred mixture of H_2O and TFA (v/v, 10 mL). The reaction mixture was stirred for 1 h at room temperature, after which the solvents were

evaporated under reduced pressure and co evaporation with toluene. Dry pyridine (5 mL) and dry acetic anhydride (5 mL) were added to the residual syrup and this reaction mixture was stirred at room temperature for 16h. The solvents were removed under reduced pressure and the resulting residue was dissolved in CH_2Cl_2 . The organic layer was washed with water and brine, dried over sodium sulfate, filtered and concentrated. The residual syrup was purified by silica flash chromatography (PE/EtOAc = 7/3) to afford **3** (892 mg, 1.78 mmol) as a yellow oil in 92% yield over 2 steps. R_f = (PE/EtOAc: 7/3 = 0.19);

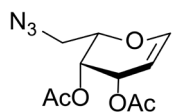
^1H NMR (400 MHz, CDCl_3) δ 7.82 – 7.70 (m, 2H), 7.37 – 7.31 (m, 2H), 5.65 (dd, J = 8.3, 1.2 Hz, 1H), 5.40 (dd, J = 2.9, 1.7 Hz, 1H), 5.32 – 5.21 (m, 1H), 5.03 (ddd, J = 10.4, 3.4, 1.2 Hz, 1H), 4.14 – 4.07 (m, 1H), 4.06 – 3.97 (m, 2H), 2.45 (s, 3H), 2.10 (s, 3H), 2.07 (s, 3H), 2.03 (s, 3H), 1.98 (s, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 170.16, 170.11, 169.67, 169.11, 145.64, 132.56, 130.32, 130.28, 128.45, 92.38, 71.63, 70.91, 67.94, 66.75, 65.63, 22.02, 21.08, 20.96, 20.84, 20.82.

3,4-Di-*O*-acetyl-6-*O*-(*p*-toluenesulfonyl)-*L*-galactal **4**



HBr (37% in AcOH, 2 mL) was added dropwise to a cooled ($-5\text{ }^\circ\text{C}$) solution of **3** (850 mg, 1.69 mmol, 1 equiv) in DCM (2 mL). The reaction was stirred for 1h at room temperature after which it was quenched by pouring into ice water. This mixture was extracted with EtOAc (three times). And the combined organic layers were washed with sat. aq. NaHCO_3 , brine, dried over sodium sulfate, filtered and concentrated. The residual crude syrup was dissolved in acetone (5 mL). Sat. aq. NaH_2PO_4 (7 mL) and Zn (dust, $<10\text{ }\mu\text{m}$, $\geq 98\%$) (1.6 gr, 24.47 mmol, 24 equiv) were added to the mixture and stirred for 3h at room temperature. The suspension was filtrated over a pad of Celite, the Celite was washed twice with acetone. The filtrate was concentrated under reduced pressure to remove excess acetone and dissolved in EtOAc. The combined organic layers were washed with NaHCO_3 , brine, dried over sodium sulfate and concentrated to obtain **4** in 77% yield over 2 steps. The product **4** was used without further purification. R_f = (PE/EtOAc:8/2 = 0.16); ^1H NMR (400 MHz, CDCl_3) δ 7.78 (d, J = 8.3 Hz, 2H), 7.39 – 7.30 (m, 2H), 6.33 (dd, J = 6.2, 1.7 Hz, 1H), 5.46 (ddd, J = 4.5, 2.2, 1.1 Hz, 1H), 5.35 (dd, J = 3.6, 1.9 Hz, 1H), 4.70 (ddd, J = 6.3, 3.1, 1.2 Hz, 1H), 4.35 – 4.04 (m, 3H), 2.44 (s, 3H), 2.03 (s, 3H), 1.99 (s, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 145.14, 129.89, 128.02, 98.85, 72.43, 67.02, 66.65, 63.69, 63.33, 21.64, 20.51, 20.51. HR MS (m/z) calcd for $\text{C}_{17}\text{H}_{20}\text{O}_8\text{S}$ [$\text{M} + \text{Na}$] $^+$, 407.0792; found, 407.0764.

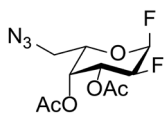
3,4-Di-*O*-acetyl-6-azido- α -*L*-galactal **5**



NaN_3 (845 mg, 13 mmol, 10 equiv) was added to a solution of **4** (500 mg, 1.30 mmol, 1 equiv) in dry DMF (5 mL) and stirred at $90\text{ }^\circ\text{C}$, with a condenser attached to the reaction vessel, until TLC showed complete conversion. H_2O and EtOAc were added. The water layer was washed three times more with EtOAc. The organic layers were concentrated under reduced pressure and purified by silica flash chromatography (PE/EtOAc = 9/1) to afford **5** (315 mg, 1.24 mmol) as a colorless oil in a 95% yield. R_f = (PE/EtOAc:8/2 = 0.5); ^1H NMR (400 MHz, CDCl_3) δ 6.48 (d, J = 6.2 Hz, 1H), 5.55 (dd, J = 2.9, 1.5 Hz, 1H), 5.43 – 5.33 (m, 1H), 4.77 – 4.73 (m, 1H), 4.25 – 4.15 (m, 1H), 3.69 – 3.27 (m, 2H), 2.14 (s, 3H), 2.03 (s, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 170.17, 170.07, 145.35, 98.93, 74.08, 64.29, 63.95,

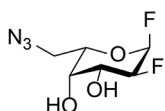
50.53, 20.76, 20.62. HR MS (m/z) calcd for $C_{10}H_{13}N_3O_5$ [$M + Na$] $^+$, 278.0792; found, 278.0763.

3,4-Di-O-acetyl-6-azido-2-fluoro- α -L-galactopyranosyl fluoride **6**



A solution of **5** (560 mg, 2.19 mmol, 1 equiv) in dry benzene (20 mL) was added dropwise (over 5 minutes) to a stirred and cooled ($-5/-10^\circ\text{C}$) solution of XeF_2 (320 mg, 1.89 mmol, 0.9 equiv) and $\text{BF}_3 \cdot \text{etherate}$ in dry ether (270 μL , 2.19 mmol, 1 equiv). The solution was allowed to warm up to room temperature and stirred for another 2h. The solution was washed with sat. aq. NaHCO_3 . After which the aqueous layers were extracted with ether. The combined organic layers were washed with H_2O , dried over sodium sulfate and concentrated. The concentrate was subsequently purified by silica flash chromatography (DCM/ether = 98/2) to afford **6** (366 mg, 1.25 mmol) as a white solid in 57% yield. R_f = (DCM/Ether:98/2=0.45); ^1H NMR (400 MHz, CDCl_3) δ 5.86 (dd, J = 53.0, 2.9 Hz, 1H), 5.51 (d, J = 3.5 Hz, 1H), 5.44 – 5.36 (m, 1H), 4.94 – 4.60 (m, 1H), 4.31 (dddd, J = 7.5, 5.3, 1.5, 0.8 Hz, 1H), 3.53 – 3.17 (m, 2H), 2.16 (s, 3H), 2.05 (s, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 105.02, 103.48, 85.63, 84.19, 70.49, 68.70, 67.71, 50.37, 20.77, 20.70. ^{19}F NMR (400 MHz, CDCl_3) δ -152.03, -211.06. HR MS (m/z) calcd for $C_{10}H_{13}F_2N_3O_5$ [$M + Na$] $^+$, 316.0715; found, 316.0714.

6-azido-2,6-dideoxy-2-fluoro- α -L-galactopyranosyl fluoride **YL209**



To a cooled solution of **6** (366 mg, 1.25 mmol, 1 equiv) in MeOH (10 mL) NaOMe was added dropwise. The reaction was stirred for 2h at room temperature. Amberlyte H $^+$ resins were added until pH = 7. The solution was filtrated and solvents were evaporated under reduced pressure. After which it was purified by silica flash chromatography (DCM/ether = 9/1) to afford YL209 as a colorless oil. Lyophilization was performed to obtain **YL209** (209 mg, mmol) as a white solid in 80% yield. R_f = (DCM/Ether:9/1=0.31); ^1H NMR (600 MHz, MeOD) δ 5.81 (dd, J = 54.3, 2.9 Hz, 1H), 4.66 (dddd, J = 48.9, 24.5, 9.9, 3.0 Hz, 1H), 4.12 (ddd, J = 8.2, 4.8, 1.3 Hz, 1H), 4.02 (ddd, J = 12.7, 9.8, 3.4 Hz, 1H), 3.96 (td, J = 3.8, 1.3 Hz, 1H), 3.64 – 3.36 (m, 2H). ^{13}C NMR (151 MHz, MeOD) δ 105.54, 88.77, 72.31, 69.78, 67.40, 50.69. ^{19}F NMR (400 MHz, MeOD) δ -153.33, -214.63. HR MS (m/z) calcd for $C_6H_9F_2N_3O_3$ [$M + Na$] $^+$, 232.0504; found, 232.0595.

Inactivation Kinetics.

The time-dependent inactivation of BfFucH by YL209 and 2FFucF was monitored measuring the residual enzyme activity over time. This was accomplished by incubation of the enzyme (0.031 mg mL $^{-1}$ BfFucH) in 100 μL of PBS buffer (+ 0.1% BSA) containing either inactivator YL209 or 2FFucF. Five inactivation concentrations (0, 5, 10, 20, 40, 80 mM) were used. The control mixtures contained the same amount of enzyme but no

inactivator. Both inactivation and control samples were incubated at room temperature and at several time intervals an aliquot of each inactivation mixture was added to a solution of the substrate, 4-nitrophenyl α -L-fucopyranoside in PBS, so that the final assay contained BfFucH at a concentration of $3.1 \mu\text{g mL}^{-1}$ and 2.2 mM substrate (5 times K_m) in PBS;⁵¹ the total volume was 100 μL . The initial rates were measured under steady-state conditions by spectrophotometric monitoring of the release of 4-nitrophenolate at 400 nm. Pseudo first order rate constants (k_{obs}) were determined by fitting the decay curve to a single exponential decay equation using Graphpad PRISM 8. The first and second order rate constants (k_{inact} and k_{inact}/K_i) were determined by fitting the pseudo first order rate constants versus the inactivator concentrations data to the equation:

$$K_{\text{obs}} = k_{\text{inact}} / (1 + K_i / [I]).$$

Reactivation Kinetics

The time-dependent reactivation of BfFucH inactivated by YLI209 was monitored measuring the residual enzyme activity over time. This was accomplished by incubation of the enzyme (0.031 mg mL^{-1} , 620 nM BfFucH) in 100 μL of PBS buffer (+ 0.1% BSA) containing inactivator YLI209 (50 mM final concentration) at 37°C for 3 h so that 75% inactivation had occurred. The control mixtures contained the same amount of enzyme but no inactivator. Excess inactivator was then removed by extensive dialysis at 4°C against large volumes of PBS using a microconcentrator (Pierce, microconcentrator 10.000 MWCO). Reactivation was monitored by addition of an aliquot, at specific time intervals, to a solution of the substrate, 4-nitrophenyl α -L-fucopyranoside in PBS, so that the final assay contained BfFucH at a concentration of $3.1 \mu\text{g mL}^{-1}$ and 2.2 mM substrate in PBS with a total volume of 100 μL . The initial rates were measured under steady-state conditions by spectrophotometric monitoring of the release of 4-nitrophenolate at 400 nm. The curve was fitted to a single exponential association equation using Graphpad PRISM 8.

Chemoselective labelling on nitrocellulose blot using alkyne-biotin.

An aliquot of the enzyme (final [BfFucH] = 100 $\mu\text{g mL}^{-1}$, 2 μM) was treated with a solution of YLI209 in PBS (final [YLI209] = 50 mM) and incubated overnight at 37 °C. The commercially available fucosidases *Thermotoga maritima* (E-FUCTM, Megazyme) and homo sapiens (E-FUCHS, Megazyme) were analyzed in the same way as the recombinant BfFucH. The samples were mixed with 3x loading buffer and the samples were loaded without heating onto a BIS-TRIS polyacrylamide gel. After a short electrophoresis of 20 minutes at 200V the samples were blotted to a nitrocellulose membrane (TransBlot Turbo Transfer System, Biorad). Allowing full protein transfer within 7 minutes to limit exposure to the basic transfer buffer. The membrane was washed for 2 minutes in PBS after which it was incubated for 2h at room temperature in the presence of 100 μM alkyne-biotin, 2.5 mM L-ascorbate and 0.5 mM Cu_2SO_4 in PBS. The click reaction was followed by 2 times 5 minutes wash with PBS containing 0.1% Tween-20 (PBS-T). The membrane was blocked using 5% skimmed milk in PBS-T with rocking for 1 h at room

temperature. The blocking solution was decanted and a solution 1% skimmed milk in PBS-T containing α biotin-HRP (1:10.000, Jackson ImmunoResearch) was added and incubated for 1h at room temperature. Membranes were washed for 3 x 10 minutes with PBS-T. Detection of membrane-bound α biotin-HRP conjugates was accomplished by chemiluminescent detection using the SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Scientific) and imaged in a Gel-Doc system (Bio-Rad).

Chemoselective labeling in solution using alkyne-biotin.

An aliquot of the enzyme (final [BfFuch] = 100 $\mu\text{g mL}^{-1}$, 2 μM) was treated with a solution of YL209 in PBS (final [YL209] = 50 mM) and incubated overnight at 37 °C. The inactivated samples were then dialysed overnight (Pierce Slide-A-lyzer Mini dialysis unit, 10.000 MWCO) at 4°C. After dialysis solutions of 1 M pH 2.0 sodium phosphate and saturated urea (one third of the volume) were added to yield a final pH of 3.5. One volume of alkyne-biotin in water (100 μM final concentration) with 2.5 mM L-ascorbate and 0.5 mM Cu_2SO_4 was added to the mixture and incubated overnight at RT. A microconcentrator (Pierce, microconcentrator 10.000 MWCO) was pre-washed with 0.1% PBS-T and centrifuged 15 minutes 15.000 x g. This was continued by the concentration of the sample in said microconcentrator to a final volume of 50 μL . The mixture was mixed with 3x laemlli buffer before loading 20 μL on a TRIS-BIS polyacrylamide gel. After electrophoresis, the gel was blotted to a nitrocellulose membrane (TransBlot Turbo Transfer System, Biorad). The membrane was blocked using 5% skimmed milk in PBS-T with rocking for 1 h at room temperature. The blocking solution was decanted and a solution 1% skimmed milk in PBS-T containing α biotin-HRP (1:10.000, Jackson ImmunoResearch) was added and incubated for 1h at room temperature. Membranes were washed for 3 x 10 minutes with PBS-T. Detection of membrane bound α biotin-HRP conjugates was accomplished by chemiluminescent detection using the SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Scientific) and imaged in a Gel-Doc system (Bio-Rad).

Acknowledgements

We would like to thank Gerlof Bosman for the expression of the recombinant protein BfFuch.

References

- 1 J. Intra, M. E. Perotti, G. Pavesi and D. Horner, *Gene*, 2007, **392**, 34–46.
- 2 K. M. Ng, J. A. Ferreyra, S. K. Higginbottom, J. B. Lynch, P. C. Kashyap, S. Gopinath, N. Naidu, B. Choudhury, B. C. Weimer, D. M. Monack and J. L. Sonnenburg, *Nature*, 2013, **502**, 96–99.
- 3 J. M. Garber, H. Nothaft, B. Pluvinage, M. Stahl, X. Bian, S. Porfirio, A. Enriquez, J. Butcher, H. Huang, J. Glushka, E. Line, J. A. Gerlt, P. Azadi, A. Stintzi, A. B. Boraston and C. M. Szymanski, *Commun. Biol.*, 2020, **3**, 1–11.
- 4 F. A. Shaikh, A. Lammerts Van Bueren, G. J. Davies and S. G. Withers, *Biochemistry*, 2013, **52**, 5857–5864.
- 5 R. A. DiCioccio, J. J. Barlow and K. L. Matta, *J. Biol. Chem.*, 1982, **257**, 714–718.
- 6 H. Sakurama, E. Tsutsumi, H. Ashida, T. Katayama, K. Yamamoto and H. Kumagai, *Biosci. Biotechnol. Biochem.*, 2012, **76**, 1022–1024.
- 7 D. A. Sela, D. Garrido, L. Lerno, S. Wu, K. Tan, H. J. Eom, A. Joachimiak, C. B. Lebrilla and D. A. Mills, *Appl. Environ. Microbiol.*, 2012, **78**, 795–803.
- 8 J. M. Pickard and A. V. Chervonsky, *J. Immunol.*, 2015, **194**, 5588–5593.
- 9 H. Wu, O. Rebello, E. H. Crost, C. D. Owen, S. Walpole, C. Bennati-Granier, D. Ndeh, S. Monaco, T. Hicks, A. Colville, P. A. Urbanowicz, M. A. Walsh, J. Angulo, D. I. R. Spencer and N. Juge, *Cell. Mol. Life Sci.*, 2020, **1**, 3.
- 10 M. T. Sorbara and E. G. Pamer, *Mucosal Immunol.*, 2019, **12**, 1.
- 11 M. Stahl, L. M. Friis, H. Nothaft, X. Liu, J. Li, C. M. Szymanski and A. Stintzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 7194–7199.
- 12 Y. M. C. A. Luijckx, N. M. C. Bleumink, J. Jiang, H. S. Overkleeft, M. M. S. M. Wösten, K. Strijbis and T. Wennekes, *Cell. Microbiol.*, 2020, **22**, e13252.
- 13 F. C. Pereira and D. Berry, *Environ. Microbiol.*, 2017, **19**, 1366–1378.
- 14 D. E. Koshland, *Biol. Rev.*, 1953, **28**, 416–436.
- 15 H. Hinou, M. Kuroguchi, H. Shimizu and S. I. Nishimura, *Biochemistry*, 2005, **44**, 11669–11675.
- 16 M. Kuroguchi, S. I. Nishimura and Y. C. Lee, *J. Biol. Chem.*, 2004, **279**, 44704–44712.
- 17 M. Nandakumar, Y. L. Hsu, J. C. Y. Lin, C. Lo, L. C. Lo and C. H. Lin, *ChemBioChem*, 2015, **16**, 1555–1559.
- 18 Y. L. Hsu, M. Nandakumar, H. Y. Lai, T. C. Chou, C. Y. Chu, C. H. Lin and L. C. Lo, *J. Org. Chem.*, 2015, **80**, 8458–8463.
- 19 T. J. Sminia, H. Zuilhof and T. Wennekes, *Carbohydr. Res.*, 2016, **435**, 121–141.
- 20 B. P. Rempel and S. G. Withers, *Glycobiology*, 2008, **18**, 570–586.
- 21 L. I. Willems, J. Jiang, K. Y. Li, M. D. Witte, W. W. Kallemeyn, T. J. N. Beenakker, S. P. Schröder, J. M. F. G. Aerts, G. A. van der Marel, J. D. C. Codée and H. S. Overkleeft, *Chem. – A Eur. J.*, 2014, **20**, 10864–10872.
- 22 J. Jiang, W. W. Kallemeyn, D. W. Wright, A. M. C. H. C. H. van den Nieuwendijk, V. Coco Rohde, E. Colomina Folch, H. van den Elst, B. I. Florea, S. Scheij, W. E. Donker-Koopman, M. Verhoek, N. Li, M. Schürmann, D. Mink, R. G. Boot, J. D. C. C. Codée, G. A. van den Marel, G. J. Davies, J. M. F. G. Aerts and H. S. Overkleeft, *Chem. Sci.*, 2015, **6**, 2782–2789.
- 23 J. Jiang, C. L. Kuo, L. Wu, C. Franke, W. W. Kallemeyn, B. I. Florea, E. Van Meel, G. A. van der Marel, J. D. C. Codée, R. G. Boot, G. J. Davies, H. S. Overkleeft and J. M. F. G. Aerts, *ACS Cent. Sci.*, 2016, **2**, 351–358.

- 24 L. Wu, J. Jiang, Y. Jin, W. W. Kallemeijn, C. L. Kuo, M. Artola, W. Dai, C. van Elk, M. van Eijk, G. A. van der Marel, J. D. C. Codée, B. I. Florea, J. M. F. G. Aerts, H. S. Overkleeft and G. J. Davies, *Nat. Chem. Biol.*, 2017, **13**, 867–873.
- 25 T. M. Gloster, R. Madsen and G. J. Davies, *Org. Biomol. Chem.*, 2007, **5**, 444–446.
- 26 S. G. Withers, K. Rupitz and I. P. Street, *J. Biol. Chem.*, 1988, **263**, 7929–7932.
- 27 S. J. Williams and S. G. Withers, *Carbohydr. Res.*, 2000, **327**, 27–46.
- 28 M. T. C. Walvoort, G. A. van der Marel, H. S. Overkleeft and J. D. C. Codée, *Chem. Sci.*, 2013, **4**, 897–906.
- 29 N. Xu, Y. Lin, S. A. Hofstadler, D. Matson, C. J. Call, B. Ren, E. M. Mosi and S. G. Withers, *64 ENZYME KINETICS AND MECHANISM [4] [4] Trapping of u-Glycosidase Intermediates*, 1998, vol. 70.
- 30 S. S. Macdonald, A. Patel, V. L. C. Larmour, C. Morgan-Lang, S. J. Hallam, B. L. Mark and S. G. Withers, *J. Biol. Chem.*, 2018, **293**, 3451–3467.
- 31 J. D. McCarter, M. J. Adam, N. G. Hartman and S. G. Withers, *Biochem. J.*, 1994, **301**, 343–348.
- 32 J. D. McCarter, M. J. Adam and S. G. Withers, *J. Label. Compd. Radiopharm.*, 1992, **31**, 1005–1009.
- 33 K. A. Stubbs, A. Scaffidi, A. W. Debowski, B. L. Mark, R. V. Stick and D. J. Vocadlo, *J. Am. Chem. Soc.*, 2008, **130**, 327–335.
- 34 D. J. Vocadlo and C. R. Bertozzi, *Angew. Chemie - Int. Ed.*, 2004, **43**, 5338–5342.
- 35 O. Hekmat, Y. W. Kim, S. J. Williams, S. He and S. G. Withers, *J. Biol. Chem.*, 2005, **280**, 35126–35135.
- 36 S. J. Williams, O. Hekmat and S. G. Withers, *ChemBioChem*, 2006, **7**, 116–124.
- 37 C. A. Tarling, S. He, G. Sulzenbacher, C. Bignon, Y. Bourne, B. Henrissat and S. G. Withers, *J. Biol. Chem.*, 2003, **278**, 47394–47399.
- 38 T. W. Liu, C. W. Ho, H. H. Huang, S. M. Chang, S. D. Popat, Y. T. Wang, M. S. Wu, Y. J. Chen and C. H. Lin, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 14581–14586.
- 39 J. D. McCarter, M. J. Adam and S. G. Withers, *Biochem. J.*, 1992, **286**, 721–727.
- 40 U. Grabowska, D. A. MacManus, K. Biggadike, M. I. Bird, S. Davies, T. Gallagher, L. D. Hall and E. N. Vulfson, *Carbohydr. Res.*, 1997, **305**, 351–361.
- 41 D. A. MacManus, U. Grabowska, K. Biggadike, M. I. Bird, S. Davies, E. N. Vulfson and T. Gallagher, *J. Chem. Soc. - Perkin Trans. 1*, 1999, 295–305.
- 42 W. Korytnyk, S. Valentekovic-Horvath and C. R. Petrie, *Tetrahedron*, 1982, **38**, 2547–2550.
- 43 C. D. Jeroen C Codée, H. S. Overkleeft, M. T. C. Walvoort, W. W. Kallemeijn, L. I. Willems, M. D. Witte, J. M. F. G. Aerts, G. A. van der Marel and J. D. C. Codeé, *Chem. Commun.*, 2012, **4848841**, 10357–10456.
- 44 H. H. Jensen and M. Bols, *Org. Lett.*, 2003, **5**, 3419–3421.
- 45 J. H. Kim, R. Resende, T. Wenekes, H. M. Chen, N. Bance, S. Buchini, A. G. Watts, P. Pilling, V. A. Streltsov, M. Petric, R. Liggins, S. Barrett, J. L. McKimm-Breschkin, M. Niikura and S. G. Withers, *Science (80-.)*, 2013, **340**, 71–75.
- 46 J. Ohata, F. Vohidov and Z. T. Ball, *Mol. Biosyst.*, 2015, **11**, 2846–2849.
- 47 J. P. Hachmann and J. W. Amshey, *Anal. Biochem.*, 2005, **342**, 237–245.
- 48 G. Sulzenbacher, C. Bignon, T. Nishimura, C. A. Tarling, S. G. Withers, B. Henrissat and Y. Bourne, *J. Biol. Chem.*, 2004, **279**, 13119–13128.

- 49 B. Doboszewski and P. Herdewijn, *Tetrahedron Lett.*, 2012, **53**, 2253–2256.
- 50 N. Lunau, K. Seelhorst, S. Kahl, K. Tscherch, C. Stacke, S. Rohn, J. Thiem, U. Hahn and C. Meier, *Chem. - A Eur. J.*, 2013, **19**, 17379–17390.
- 51 T. I. Tsai, S. T. Li, C. P. Liu, K. Y. Chen, S. S. Shivatare, C. W. Lin, S. F. Liao, C. W. Lin, T. L. Hsu, Y. T. Wu, M. H. Tsai, M. Y. Lai, N. H. Lin, C. Y. Wu and C. H. Wong, *ACS Chem. Biol.*, 2017, **12**, 63–72.

Supplementary Figure

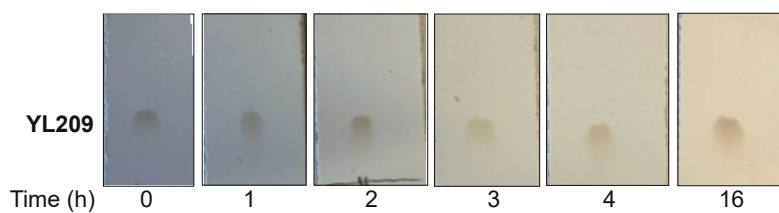
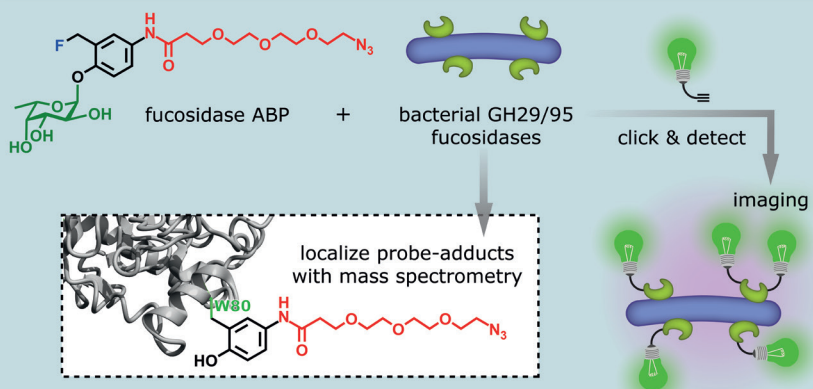


Figure S1 | TLC stability of **YL209** in PBS at room temperature. Followed over time, run in Hexane : diethyl ether (3:1), visualized with 5% sulfuric acid in ethanol followed by heating.



Chapter 5

Detection of bacterial α -L-fucosidases with an *ortho*-quinone methide-based probe and mapping of the probe-protein adducts

Yvette M.C.A. Luijkx, Anniëk J. Henselijn, Gerlof P. Bosman, Dario A.T. Cramer, Koen C.A.P. Giesbers, Esther M. van 't Veld, Geert-Jan P.H. Boons, Albert J.R. Heck, Karli R. Reiding, Karin Strijbis, Tom Wennekes

Abstract

An *ortho*-quinone methide-based probe with an azide mini-tag that selectively labels both retaining and inverting bacterial α -L-fucosidases is reported. Mass spectrometry-based intact protein and sequence analysis of a probe-labelled bacterial fucosidase revealed almost exclusive single labelling at two specific tryptophan residues outside of the active site. Furthermore, the probe could detect and image extracellular fucosidase activity on the surface of live bacteria.

Introduction

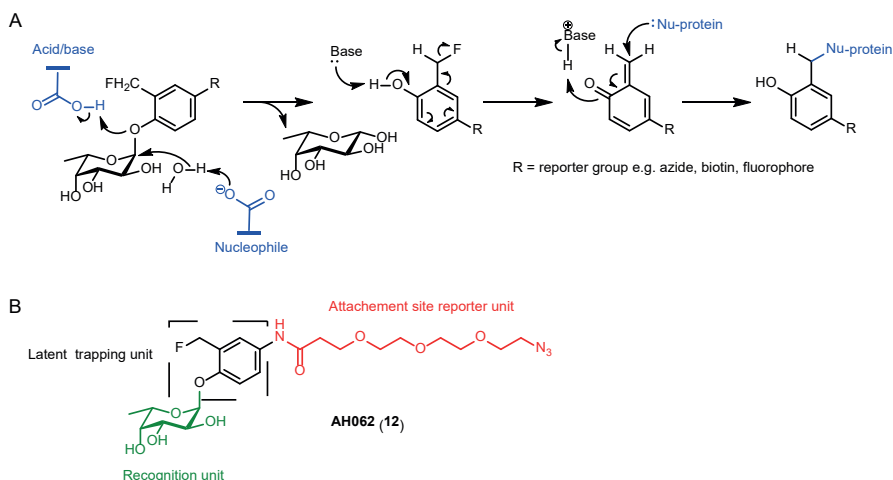
α -L-Fucosidases are enzymes capable of catalyzing the hydrolytic removal of terminal L-fucosyl residues from glycoconjugates. It is known that an abnormal increase of α -L-fucosidase activity in humans is associated with several pathological conditions.^{1–4} Furthermore, bacterial α -L-fucosidases are key enzymes for the degradation and metabolism of intestinal mucin *O*-glycans by gut microbes. This crucial family of enzymes thereby contributes to the composition of the gut microbiota and influences our health and disease.⁵ We and others recently reported that fucosidases of two *Bacteroides* species from the human gut microbiota induced upregulation of growth and invasive properties of pathogenic *Campylobacter jejuni* strains.^{6,7} A growing number of studies are implicating bacterial fucosidases in host-microbe interplay in the intestine, microbiota cross feeding, and colonization resistance.^{3,5,8–10}

α -L-Fucosidases employ one of two mechanisms to catalyze the hydrolysis of terminal α -fucosidic linkages and these result in either overall retention or inversion of the anomeric configuration. Fucosidases employing a double-displacement (Koshland) retaining mechanism belong to the glycoside hydrolase family 29 (GH29), whereas those employing an inverting mechanism belong to the glycoside hydrolase family 95 (GH95).^{11,12} Given their importance and the vast amounts of fucosidases in the human-gut microbiome, several attempts have been made to develop activity-based probes (ABPs) to elucidate their individual functions. However, the majority of ABPs developed against fucosidases exploit the double-replacement retaining mechanism of GH29 fucosidases and are thus incapable of targeting GH95 fucosidases.^{13–16} Thereby highlighting the need to expand the molecular toolbox to study all known and still to be discovered bacterial α -L-fucosidases.

A type of ABPs that would allow the labelling of both GH29 as well as GH95 fucosidases are the quinone methide-based probes.^{17,18} Upon α -L-fucosidase cleavage of the glycosidic bond in such a probe, the liberated *ortho*-fluoromethylphenolate aglycone undergoes rapid 1,4-elimination through expulsion of a fluoride ion to generate a highly electrophilic *ortho*-quinone methide. This reactive intermediate quickly alkylates a nearby nucleophile, which results in the formation of a covalently probe-labeled enzyme adduct (Scheme 1A). This type of probe has already been shown capable of labelling β -galactosidases, neuraminidases of *V. cholera* and human GH29 α -L-fucosidase enzymes.^{4,17,19,20} One of the challenges in the application of quinone methide-based probes is the possibility for off-target labelling of nucleophiles in nearby biomolecules due to diffusion of the *ortho*-quinone methide with its relatively long lifetime. However, this property also makes the probe an interesting candidate for imaging

of enzyme activity as the probe does not react with the active site which would lead to inactivation of the enzyme. Instead the possibility for repeated hydrolysis of new copies of the probe could lead to accumulation of multiple adducts nearby the active site. This strategy with multiple adducts has previously been applied for signal amplification in imaging studies.^{21,22}

A quinone methide-based probe to detect fucosidase activity has previously been developed by Hsu *et al.*¹⁹ However, the application of this probe is limited because of the bulky hydrophobic reporter group that might prevent the probe from accessing the active site of some α -L-fucosidases in gut microbiota.²³ Instead, the installation of a mini-tag, such as an azide or an alkyne, could overcome this limitation and make the probe more flexible in the detection of labelling. Thus, to enable effective *in vivo* imaging of both GH29 and GH95 bacterial fucosidases we developed a more compact and flexible version of the Hsu *et al.* probe by the incorporation of an azide mini-tag. Conjugation of a reporter group to the probe following labelling of the enzyme target was accomplished by the azide mini-tag reacting via Huisgen's 1,3-dipolar cycloaddition with a variety of alkyne-bearing reporter groups (Scheme 1B).¹⁹



Scheme 1 | A) Activity-based covalent labelling of α -L-fucosidases with quinone methide functionalized ABP. B) Design of probe **AH062 (12)**.

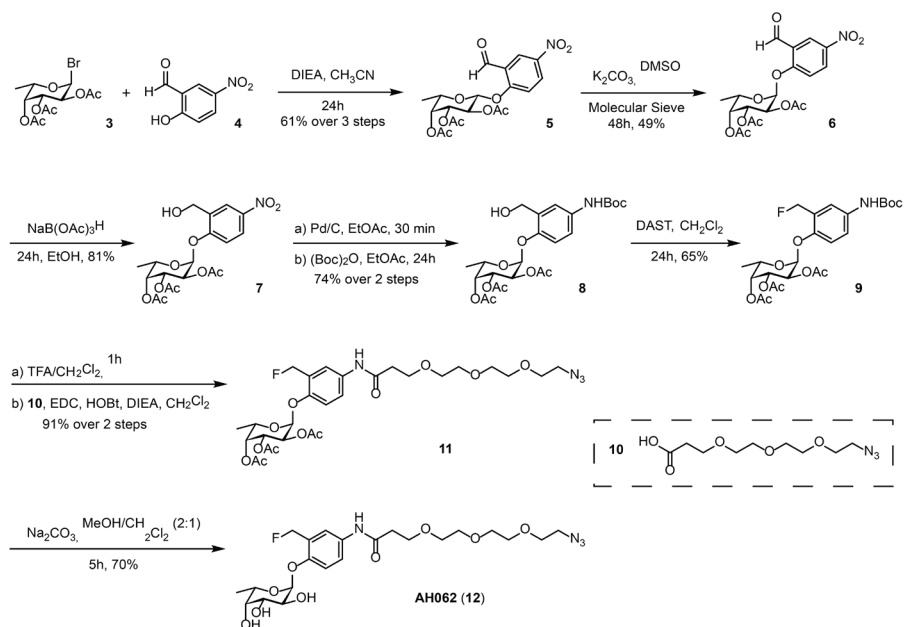
Results and discussion

We selected a synthetic strategy previously published by Hsu *et al.* to develop a synthetic route for our activity-based probe **AH062 (12)**, beginning with the commercially available L-fucose. Peracetylation of L-fucose and subsequently treatment with hydrogen bromide in acetic acid provided protected fucosyl bromide **3**.²⁴ Without further

purification, donor **3** was coupled with the acceptor, 2-hydroxy-5-nitrobenzaldehyde **4**, under classical Koenigs-Knorr conditions to provide the 1,2-*trans* glycosylated product **5** in 61% yield over 3 steps (Scheme 2).¹⁹

Subsequently, compound **5** underwent a base-promoted anomeric epimerization upon treatment with K_2CO_3 in DMSO under dry conditions to afford the desired 1,2-*cis* product **6** in 49% yield, with concomitant recovery of **5** (25%).¹⁹ Next, the fluoromethylaryl group of compound **6** was reduced with $NaBH(OAc)_3$ to the benzylic alcohol **7** in 81% yield. The nitro group of **7** was then reduced to the intermediate arylamine by catalytic hydrogenation that was immediately Boc-protected to furnish compound **8** in 74% yield over 2 steps. Treatment of **8** with Diethylaminosulfur trifluoride (DAST) in CH_2Cl_2 provided the benzylic fluoride **9** in 65% yield. Compound **10** was prepared beforehand from commercially available $H_2N-PEG_3-CH_2CH_2COOH$ (PurePEG) through conversion of the amine group to the corresponding azide with diazotransfer reagent imidazole-1-sulfonyl azide hydrochloride.²⁵ After removal of the Boc-protecting group in **8** upon treatment with TFA, amide formation with carboxylic acid **10** under standard carbodiimide coupling conditions (1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide / 1-hydroxybenzotriazole (EDC/HOBt)) gave compound **11** in high yield (91% over 2 steps). Finally, compound **11** was deacetylated under mild conditions (Na_2CO_3 in MeOH) to provide the target probe **AH062 (12)** in 70% yield.

To determine the labelling efficiency of **AH062**, we incubated recombinant *B. fragilis* fucosidase (BfFucH) in various concentrations (4 – 0.25 μg corresponding to 60-960 nM) with **AH062** (350 μM) for 2h at 37 °C. The resulting mixture was analyzed by SDS-PAGE followed by Western blotting. Detection was achieved by a Cu(I)-catalyzed click reaction on the blot surface with alkyne-biotin followed by incubation with anti-biotin HRP-linked antibody and detection using enhanced chemiluminescence (ECL) substrate. The probe was capable of detecting the BfFucH enzyme at concentrations as low as 0.25 μg (60 nM) (Figure 1A). **AH062** was also capable of distinguishing the active from the inhibited enzyme (prepared by preincubation with the competitive inhibitor FNJ). Generally, a concern with quinone-methide based probes is their labelling selectivity, due to diffusion of the *ortho*-quinone methide trapping unit away from the targeted enzyme. To determine labelling selectivity, we examined the ability of **AH062** to specifically target fucosidases in the presence of a second protein. BfFucH samples were spiked with equimolar concentration transferrin (TF) and incubated with the probe. TF is a glycoprotein of 79.5 kDa and can therefore be clearly separated on gel from the ~50 kDa BfFucH. **AH062** showed clear signals for BfFucH, but no chemiluminescent signal was detected for TF (Figure 1B). These results suggest that the **AH062** is specifically reactive in or near the active site of the fucosidase. Furthermore, treatment of TF with **AH062** in the absence of active BfFucH did not give any detectable signal.



Scheme 2 | Synthesis of probe **AH062** (12).

To assess the ability of **AH062** to visualize catalytically active inverting GH95 fucosidases, we expressed the known GH95 fucosidase AfcA from *Bifidobacterium bifidum*. AfcA was incubated with **AH062** at five different concentrations (8 – 3.7 μg) and the probe detected the enzyme in a concentration-dependent manner (Figure 1C). **AH062** labelled as little as 4.8 μg of AfcA (Figure 1C). To our knowledge, this is the first example of an ABP capable of labelling of an inverting fucosidase (GH95).

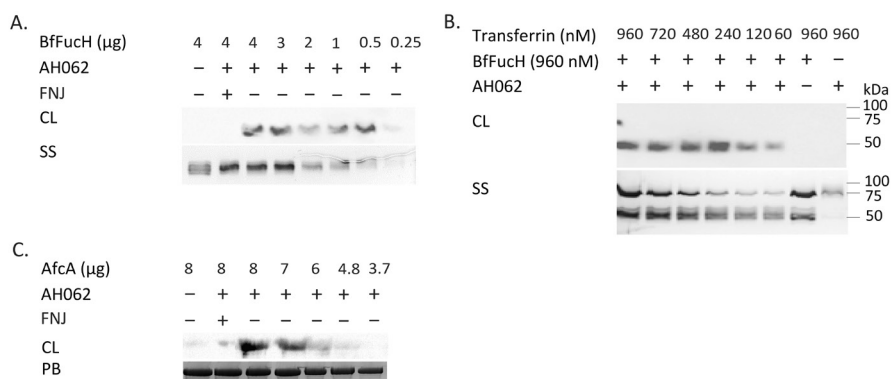


Figure 1 | *In vitro* labelling of α -L-fucosidases with **AH062**. A) Limit of detection by **AH062**. Various concentrations of BfFucH were incubated with 350 μ M of **AH062** in PBS for 2 h at room temperature. Silver staining was included as loading control. B) Selective labelling of BfFucH (50 kDa) with **AH062** in the presence of decreasing concentrations of TF (80 kDa). Silver staining was included as loading control. C) Different concentrations of AfcA were incubated with 350 μ M of **AH062** in PBS for 2 h at room temperature. PageBlue was included as loading control. CL, chemiluminescence, SS, silver staining. PB, PageBlue.

We next performed Mass spectrometric studies to determine the attachment sites of the novel quinone-methide probe onto the fucosidase enzyme. After incubation of **AH062** for 2h at 37 °C with BfFucH, intact LC-MS profiles from labelled and untreated samples were compared. The spectrum of labelled proteins showed approximately 6.1% of the fucosidases had a mass shift of 350 Da that fit with a single probe-protein adduct. Additionally < 0.5% of the fucosidases had a mass shift that corresponds with two probe-adducts on BfFucH (Figure S1).

To identify the specific amino acids on the fucosidase surface that were being labeled by the probe, both labelled and unlabelled samples were digested with pepsin at low pH followed by LC-MS/MS analysis. Unfortunately, the resulting BfFucH samples did not yield good coverage for the adduct-labeled peptide fragments. Therefore, we repeated this analysis with the well-characterized bacterial GH29 TmFuc fucosidase for bottom-up analysis. We were now able to identify two unique peptide fragments that were labelled by **AH062**. To determine the exact site of the modification, tandem mass analysis was performed, whereby the y_n and b_n were assigned based on the annotation proposed by Roepstorff and Fohlman (Figure 2A and Figure S2). This data revealed that the quinone methide generated from **AH062** had either labeled tryptophan W80 and/or W105. In Figure 2B the modification sites are mapped on the known crystal structure of TmFuc. The modifications on W80 and W105 are positioned outside the fucosidase active site, separated respectively 27 Å and 28 Å from active site nucleophile D224. Although quinone methides have been reported to react with various nucleophiles, the one generated from **AH062** shows a high preference for the surface tryptophans in TmFuc.

Previous literature has shown high selectivity of a very similar quinone methide for tryptophan residues in proteins.²⁶

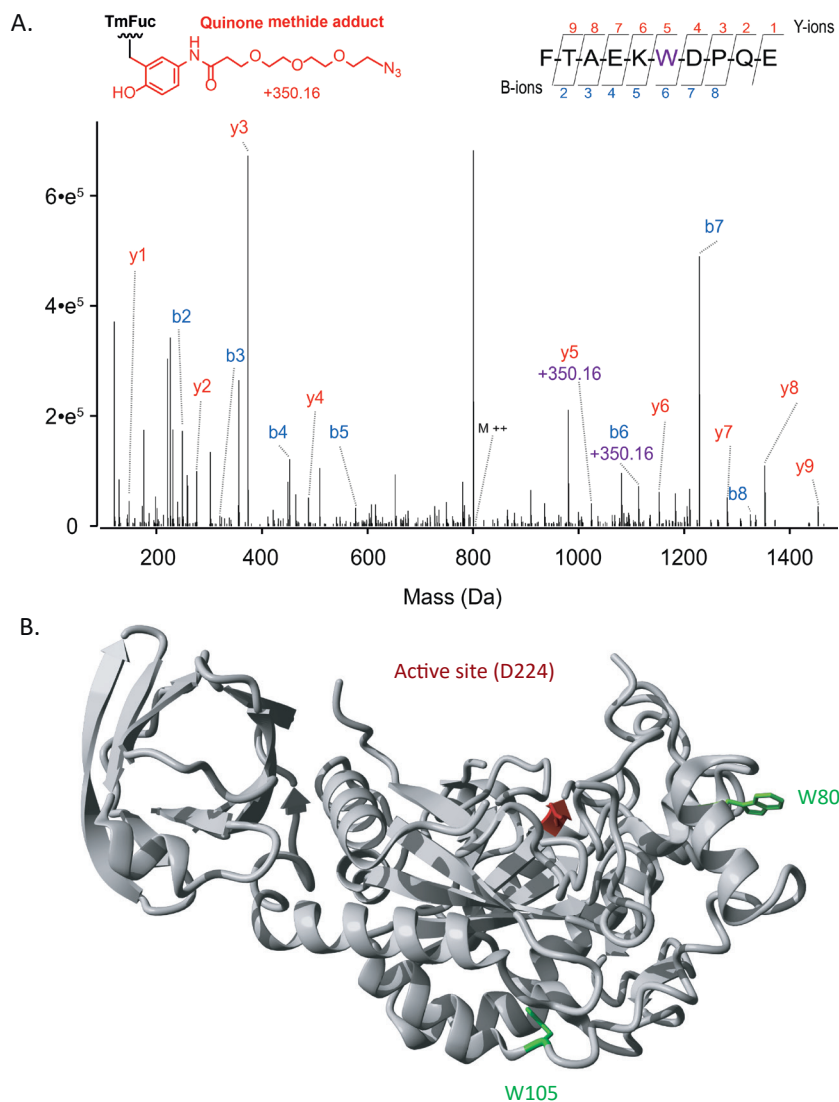


Figure 2 | A) Tandem MS spectrum of a by pepsin generated peptide of TmFuc, revealing its labelling at W105. The annotated y and b ion series are shown above and below the peptide sequence, respectively. The y5 and b6 carry the mass increment of the probe (+350.16 Da), indicating the modification to be on the W105 residue. B) Structure of TmFuc with its active-site catalytic acid D224 annotated in red and residues modified by quinone methide probe **AH062** annotated in green.

To further assess the utility of this probe for imaging bacterial fucosidase activity, we used **AH062** to stain fucosidases associated with the cell wall of *B. fragilis*. Overnight cultures of *B. fragilis* were incubated for 2h at 37 °C with **AH062**. As a negative control, bacterial samples were taken along that were preincubated with 100 μ M competitive fucosidase inhibitor FNJ followed by incubation with the probe. To visualize the probe, the bacteria were clicked by copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) to an Alexa488 dye and the bacterial cell membranes were stained with CTY to facilitate locating the bacteria (Figure 3A). Incubation with **AH062** resulted in highly specific fluorescence labelling that was not observed after preincubation with FNJ. The **AH062** pattern of labelling showed a similar localization on the outer membrane compared to the CTY cell membrane staining. By applying high resolution Structured illumination microscopy (SIM), we confirmed that **AH062** stained the *B. fragilis* cell wall or cell membrane (Figure 3B). We next performed FACS analysis of **AH062**-labeled bacteria and observed a shift in the whole peak indicating that the majority of the bacteria were fluorescently labeled by covalent attachment of the activated quinone methide. Again, labelling with **AH062** could be blocked by pre incubation with FNJ showing the selectivity of the probe (Figure S3).

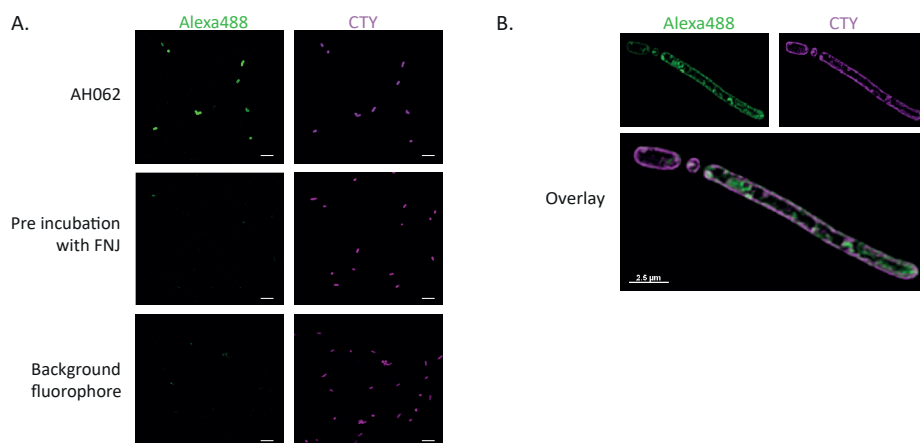


Figure 3 | Immunofluorescence confocal microscopy imaging of **AH062**-labelled *B. fragilis*. A) Selective labelling of active fucosidases of *B. fragilis* with **AH062** (green) and overall bacterial cell membrane stain with CTY (purple). Scale bars are 5 μ m. B) High-resolution SIM imaging of membrane labelling by **AH062** and CTY in a single *B. fragilis* bacterium. SIM Images show maximum intensity projections.

Conclusion

In conclusion, we report a novel fucosidase probe (**AH062**) that is capable of labelling both GH29 and GH95 bacterial fucosidases, making it the first probe that targets inverting fucosidases. The quinone-methide probe mainly labels on specific tryptophan

residues outside of the active site of the fucosidase enzyme with single protein-probe adducts dominating. The probe was successfully used for the detection of fucosidase-positive bacteria by high resolution confocal microscopy and flow cytometry. For further application of this probe in complex co-culture systems, the likely occurrence of quinone-methide labelling adducts on off-target proteins in co-cultures needs to be determined and minimized. In the future, we envision applying this novel probe for labelling, identification and characterization of important bacterial GH29 and GH95 fucosidases at the intestinal host-microbe interface.

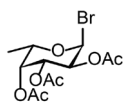
Experimental

General methods and materials

All reagents and starting materials were obtained from commercial suppliers and were used without further purification. MeCN and CH_2Cl_2 were dried and purified using an MBraun MB SPS 800 prior to use. Pyridine and DMSO were dried for 24h over pre-activated (5 min, $\sim 300^\circ\text{C}$) 4Å molecular sieves prior to use. Glassware for anhydrous reactions was flame-dried and cooled under a nitrogen atmosphere immediately prior to use. Analytical TLC was performed on glass-backed TLC plates pre-coated with silica gel (60G, F_{254}). Compound **10** was visualised by staining with molybdenum, ninhydrin and 10% PPh_3 in CH_2Cl_2 followed by ninhydrin. All other compounds were visualised under UV light and/or by staining with molybdenum and 5% sulfuric acid in EtOH. Column chromatography was performed with silica gel (230-400 mesh). ^1H NMR, ^{13}C NMR and ^{19}F NMR spectra were recorded on either a Bruker Avance 600 MHz NMR spectrometer or 400 MHz spectrometer in the designated deuterated solvents (CDCl_3 or CD_3OD). ^1H and ^{13}C NMR peak assignments are made based on ^1H - ^1H COSY and ^1H - ^{13}C HSQC experiments and, where possible, compared to previously reported data.¹⁹ Chemical shifts (δ) are listed in ppm downfield from TMS using TMS as an internal reference. Coupling constants are reported in Hz. The following abbreviations were used to explain the multiplicities: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, b = broad. Expression and purification of the recombinant proteins BfFucH and AfcA were based on previous publications.^{11,27} Recombinant TmFuc and tissue-derived transferrin (TF) both of sufficient quality were commercially available via respectively Megazyme (9037-65-4) and Calbiochem (616397).

Synthesis

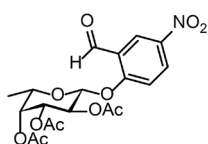
(2S,3S,4R,5R,6S)-2-bromo-6-methyltetrahydro-2H-pyran-3,4,5-triyl triacetate **3**



L-Fucose (1 g, 6.1 mmol) was dissolved in pyridine (6 mL) and Ac₂O (6 mL) and the reaction mixture was stirred for 5h at room temperature. Solvents were removed under reduced pressure and the residue obtained was dissolved in 20 mL DCM. The solution was washed with 1M HCl (3 x 20 mL), H₂O (3 x 20 mL), and brine (20 mL), dried over anhydrous Na₂SO₄, filtered, and concentrated under reduced pressure. Purification of per-acetylated L-fucose by silica gel column chromatography (eluent: 80% PetEt/EtOAc) provided the desired product in 99% yield as a colourless oil. *R*_f=0.45 (PetEt/EtOAc = 2/1).

To an ice-cooled solution of tetra-*O*-acetyl- α -L-fucopyranose in anhydrous CH₂Cl₂ (5 mL) was added dropwise HBr in AcOH (33 wt%, commercially available). The reaction was stirred for 30 min at room temperature, then poured into ice water (100 mL). The solution was extracted with CHCl₃ (100 mL) and washed with ice water (100 mL x 3), NaHCO₃ (50 mL x 3), and brine (100 mL). The organic phase was dried over Na₂SO₄, filtered, and concentrated under reduced pressure. The crude product **3** (a colourless oil) was then used in the next reaction step without further purification. *R*_f = 0.58 (PetEt/EtOAc = 2/1).

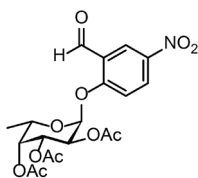
(2R,3S,4R,5R,6S)-2-(2-formyl-4-nitrophenoxy)-6-methyltetrahydro-2H-pyran-3,4,5-triyl triacetate **5**



DIPEA (2 mL, 12.33 mmol) was slowly added to an ice-cooled solution of crude bromide **3** (1.452 g, 4.11 mmol) and 2-hydroxy-5-nitrobenzaldehyde **4** (756 mg, 4.52 mmol) in 8 mL of anhydrous MeCN. The reaction mixture was stirred at room temperature for 24h under argon. After completion of the reaction, the mixture was concentrated under reduced pressure. The residue was dissolved in EtOAc (50 mL) and the resulting solution was washed with 5% citric acid (3 x 25 mL), 5% NaHCO₃ (3 x 25 mL), H₂O (3 x 25 mL), and brine (25 mL). The organic phase was dried over anhydrous Na₂SO₄, filtered, and concentrated under reduced pressure which yielded a dark orange powder. Purification by silica gel column chromatography (eluent: 70% PetEt/EtOAc) provided the desired glycosylated product **5** (1.21 g, 2.75 mmol) as a white solid in 61% yield over 3 steps. *R*_f = 0.40 (PetEt/EtOAc = 6/4). ¹H NMR (CDCl₃, 400 MHz): δ 10.35 (s, 1H), 8.72 (d, *J* = 2.9 Hz, 1H), 8.43 (dd, *J* = 9.1, 2.8 Hz, 1H), 7.23 (s, 1H), 5.59 (dd, *J* = 10.5, 7.9 Hz, 1H), 5.36 (d, *J* = 3.4 Hz, 1H), 5.27 (d, *J* = 7.8 Hz, 1H), 5.17 (dd, *J* = 10.5, 3.4 Hz, 1H), 4.08 (q, *J* = 6.3 Hz, 1H), 2.22 (s, 3H), 2.06 (s, 3H), 2.03 (s, 3H), 1.31 (d, *J* = 6.4 Hz, 3H).

¹H spectrum of this known compound was identical to previously reported spectrum.¹⁹

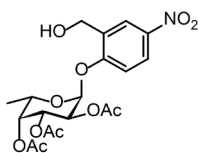
(2S,3S,4R,5R,6S)-2-(2-formyl-4-nitrophenoxy)-6-methyltetrahydro-2H-pyran-3,4,5-triyl triacetate **6**



To a mixture of compound **5** (1.39 g, 3.16 mmol), a small amount of activated 3 Å molecular sieves (~3 beads), and K_2CO_3 (4.36 g, 31.6 mmol) was added dropwise 11 mL of anhydrous DMSO. The mixture was stirred for 48h at room temperature under argon, diluted with 50 mL $CHCl_3$, and washed with 5% $NaHCO_3$ (3 x 25 mL). The combined aqueous layers were re-extracted with 25 mL $CHCl_3$ and the combined organic phases were washed with 5% citric acid (2 x 25 mL) and brine (25 mL). The organic phase was dried over Na_2SO_4 , filtered, and concentrated under reduced pressure which yielded a bright orange oil. Purification by silica gel column chromatography (eluent: 70% PetEt/EtOAc + 1% AcOH) provided desired epimer **6** (0.68 g, 1.35 mmol) in 49% yield as a pale yellow solid and the recovered yield for **5** (0.39 g, 0.88 mmol) was 28%. R_f = 0.73 (PetEt/EtOAc = 1/1). 1H NMR ($CDCl_3$, 400 MHz): δ 10.53 (s, J = 1.5 Hz, 1H), 8.73 (d, J = 2.8 Hz, 1H), 8.44 (dd, J = 9.2, 2.9 Hz, 1H), 7.43 (d, J = 9.2 Hz, 1H), 5.94 (d, J = 3.5 Hz, 1H), 5.53 (dd, J = 11.0, 3.2 Hz, 1H), 5.42 (d, J = 3.2 Hz, 1H), 5.38 (dd, J = 10.9, 3.6 Hz, 1H), 4.27 (q, J = 6.4 Hz, 1H), 2.22 (s, 3H), 2.06 (s, 3H), 2.05 (s, 3H), 1.19 (d, J = 6.5 Hz, 3H).

1H spectrum of this known compound was identical to previously reported spectrum.¹⁹

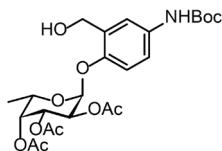
(2S,3S,4R,5R,6S)-2-(2-(hydroxymethyl)-4-nitrophenoxy)-6-methyltetrahydro-2H-pyran-3,4,5-triyl triacetate **7**



To a solution of compound **6** (388 mg, 0.88 mmol) in EtOH was added $NaBH(OAc)_3$ (655 mg, 3.1 mmol). The reaction mixture was stirred for 24h under nitrogen, diluted with EtOAc (100 mL), and washed with H_2O (3 x 50 mL) and brine (50 mL). The organic phase was dried over Na_2SO_4 , filtered, and concentrated under reduced pressure. Purification by silica gel column chromatography (eluent: 80% PetEt/EtOAc) provided desired product **7** (311 mg, 0.71 mmol) in 81% yield as a white solid. R_f = 0.58 (PetEt/EtOAc = 1/1). 1H NMR ($CDCl_3$, 400 MHz): δ 8.29 (d, J = 2.4 Hz, 1H), 8.17 (dd, J = 9.0, 2.5 Hz, 1H), 7.26 (d, J = 9.1 Hz, 1H), 5.81 (d, J = 3.6 Hz, 1H), 5.52 (dd, J = 10.9, 3.2 Hz, 1H), 5.43 – 5.35 (m, 2H), 4.89 (d, J = 13.8 Hz, 1H), 4.67 (dd, J = 13.8 Hz, 6.4 Hz, 1H), 4.25 (q, J = 6.5 Hz, 1H), 2.49 (b, 1H), 2.22 (s, 3H), 2.09 (s, 3H), 2.07 (s, 3H), 2.05 (s, 3H), 1.17 (d, J = 6.5 Hz, 3H).

1H spectrum of this known compound was identical to previously reported spectrum.¹⁹

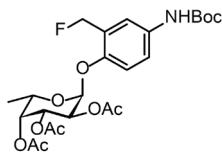
(2S,3S,4R,5R,6S)-2-(4-(tert-butoxycarbonylamino)-2-(hydroxymethyl)phenoxy)-6-methyltetrahydro-2H-pyran-3,4,5-triyl triacetate 8



To a solution of compound **7** (311 mg, 0.71 mmol) in EtOAc was added Pd/C catalyst (10 wt. % Pd, 62 mg). The reaction mixture was purged and filled with H₂, then stirred for 30 min at room temperature. Thereafter, Boc₂O (162 mg, 0.74 mmol) was added to the reactor and the solution was stirred for 24h at room temperature. More Boc₂O (1 eq) was added and the solution was again stirred for 24h. The reaction mixture was filtered over a fritted glass funnel packed with Celite and concentrated under reduced pressure. The residue was diluted with EtOAc and washed with H₂O (3 x 25 mL) and brine (25 mL). The organic phase was dried over Na₂SO₄, filtered, and concentrated under reduced pressure. Purification by silica gel column chromatography (eluent: 70% PetEt/EtOAc + 1% AcOH) provided the desired product **8** (265 mg, 0.52 mmol) in 74% yield as white solid. *R*_f = 0.71 (PetEt/EtOAc = 4/6). ¹H NMR (CDCl₃, 400 MHz): δ 7.30 (s, 1H), 7.23 (d, *J* = 8.8 Hz, 1H), 7.10 (d, *J* = 8.8 Hz, 1H), 6.38 (s, 1H), 5.58 (d, *J* = 3.6 Hz, 1H), 5.49 (dd, *J* = 10.9, 3.1 Hz, 1H), 5.42 – 5.33 (m, 2H), 4.85 (dd, *J* = 12.9, 4.7 Hz, 1H), 4.49 (dd, *J* = 12.9, 8.5 Hz, 1H), 4.33 (q, *J* = 6.4 Hz, 1H), 2.19 (s, 3H), 2.07 (s, 3H), 2.01 (s, 3H), 1.50 (s, 9H), 1.17 (d, *J* = 6.5 Hz, 3H).

¹H spectrum of this known compound was identical to previously reported spectrum.¹⁹

(2S,3S,4R,5R,6S)-2-(4-(tert-butoxycarbonylamino)-2-(fluoromethyl)phenoxy)-6-methyltetrahydro-2H-pyran-3,4,5-triyl triacetate 9

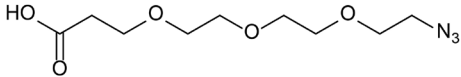


To an ice-cooled solution of compound **8** (204 mg, 0.40 mmol) in 2.2 mL of anhydrous CH₂Cl₂ was added a solution of DAST (79 μL, 0.6 mmol) in 2.2 mL of anhydrous CH₂Cl₂. The reaction mixture was allowed to warm up to room temperature and was stirred for 3.5h under argon atmosphere. The reaction was quenched by adding a small amount of silica and 10 drops of MeOH, followed by stirring for 10 min. Then, the reaction mixture was filtered over a fritted glass funnel and concentrated under reduced pressure. Purification by silica gel column chromatography (eluent: 80% PetEt/EtOAc) provided the desired product **9** (133 mg, 0.26 mmol) in 65% yield as a pale yellow oil. *R*_f = 0.78 (PetEt/EtOAc = 1/1). ¹H NMR (CDCl₃, 600 MHz): δ 7.42 (s, *J* = 15.1 Hz, 1H), 7.30 (d, *J* = 7.7 Hz, 1H), 7.10 (d, *J* = 8.8 Hz, 1H), 6.48 (s, 1H), 5.61 (d, *J* = 3.7 Hz, 1H), 5.55 – 5.32 (m, 4H), 5.28 (dd, *J* = 10.9, 3.7 Hz, 1H), 4.31 (q, *J* = 6.5 Hz, 1H), 2.20 (s, 3H), 2.06 (s, 3H), 2.03 (s, 3H), 1.51 (s, 9H), 1.16 (d, *J* = 6.5 Hz, 3H). ¹³C NMR (CDCl₃, 151 MHz): δ 170.59 (C), 170.47 (C), 170.13 (C), 152.94 (C), 150.43 (d, *J* = 4.3 Hz, C), 133.50 (C), 126.53 (d, *J* = 16.8 Hz, C), 120.65 (CH), 120.05 (CH), 115.91 (CH), 96.19 (CH), 80.36 (C), 79.78 (d, *J* = 165.8 Hz, CH₂F), 70.85 (CH), 67.97 (CH), 67.76 (CH), 65.50 (CH), 28.32 (CH₃), 20.76 (CH₃), 20.70 (CH₃), 20.66 (CH₃), 15.90 (CH₃). ¹⁹F NMR

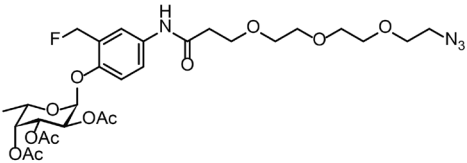
(376 MHz, CDCl_3): δ -212.13 (t, J = 52.1 Hz).

^1H and ^{13}C spectrum of this known compound was identical to previously reported spectrum.¹⁹

3-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)propanoic acid **10**

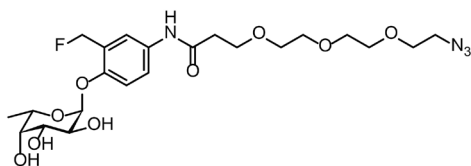
 Imidazole-1-sulfonyl azide hydrochloride (231 mg, 1.1 mmol) was added to a solution of acid-PEG3-amine (204 mg, 0.92 mmol), K_2CO_3 (280 mg, 2.02 mmol), and $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ (2.3 mg, 0.01 mmol) in 5 mL MeOH. The reaction mixture was stirred for 6h at room temperature and concentrated under reduced pressure without heating. The residue was diluted with 15 mL H_2O and acidified by dropwise addition of 1M HCl to pH 4. The mixture was extracted with EtOAc (3 x 10 mL) and the combined organic layers are dried over Na_2SO_4 , filtered and concentrated under reduced pressure. Purification by silica gel column chromatography (eluent: 60% PetEt/EtOAc) provided the desired product **10** (196 mg, 0.79 mmol) in 86% yield as a colourless oil. R_f = 0.18 (PetEt/EtOAc = 1/1). ^1H NMR (600 MHz, CDCl_3) δ 3.77 (t, J = 6.3 Hz, 2H), 3.72 – 3.59 (m, 10H), 3.40 (t, J = 4.9 Hz, 2H), 2.65 (t, J = 6.3 Hz, 2H). ^{13}C NMR (151 MHz, CDCl_3) δ 176.52 (CO_2H), 70.66 (CH_2), 70.63 (CH_2), 70.45 (CH_2), 70.41 (CH_2), 70.02 (CH_2), 66.30 (CH_2), 50.66 (CH_2N_3), 34.82 (CH_2). HRMS (ESI-quadrupole) calculated for $\text{C}_9\text{H}_{17}\text{N}_3\text{O}_5$ ($M - \text{H}$): 246.1095; found: 246.1096.

(2S,3S,4R,5R,6S)-2-(4-(3-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)propanamido)-2-(fluoromethyl)phenoxy)-6-methyltetrahydro-2H-pyran-3,4,5-triyl triacetate **11**

 Compound **9** (156 mg, 0.304 mmol) was dissolved in 1 mL of 20% TFA/ CH_2Cl_2 and the solution was stirred for 1.5h at room temperature. The volatiles were concentrated under reduced pressure and the flask was kept under high vacuum for 3h to remove residual TFA. The residue obtained was dissolved in 2.3 mL of CH_2Cl_2 and to the solution was added compound **10** (83 mg, 0.336 mmol), HOBT (21 mg, 0.152 mmol), DIEA (98 mg, 0.758 mmol), and EDC (117 mg, 0.610 mmol). The reaction mixture was stirred for 24h under nitrogen, diluted with EtOAc and washed with 5% citric acid (3 x 25 mL), 5% NaHCO_3 (3 x 25 mL), H_2O (3 x 25 mL), and brine (25 mL). The organic phase was dried over anhydrous Na_2SO_4 , filtered, and concentrated under reduced pressure. Purification by silica gel column chromatography (60% PetEt/EtOAc) provided the desired product **11** (178 mg, 0.277 mmol) in 91% yield as a pale yellow oil. R_f = 0.17 (PetEt/EtOAc = 1/1). ^1H NMR (600 MHz, CDCl_3): δ 8.65 (s, 1H), 7.59 (d, J = 8.1 Hz, 1H), 7.52 (s, 1H), 7.12 (d, J = 8.8 Hz, 1H), 5.64 (d, J = 3.2 Hz, 1H), 5.57 – 5.34 (m, 4H), 5.28 (dd, J = 10.9, 3.4 Hz, 1H), 4.31 (q, J = 6.4 Hz, 1H),

3.84 – 3.79 (m, 2H), 3.75 – 3.59 (m, 10H), 3.38 – 3.28 (m, 2H), 2.65 (m, 2H), 2.36 (s, 3H), 2.20 (s, 3H), 2.03 (s, 3H), 1.16 (d, $J = 6.4$ Hz, 3H). ^{13}C NMR (151 MHz, CDCl_3): δ 170.58 (C), 170.47 (C), 170.13 (d, $J = 4.5$ Hz, C), 133.48 (C), 122.32 (CH), 121.62 (d, $J = 7.4$ Hz, CH), 115.69 (CH), 96.14 (CH), 79.78 (d, 165.8 Hz, CH_2F), 70.85 (CH), 70.67 (CH_2), 70.52 (CH_2), 70.37 (CH_2), 70.25 (CH_2), 70.01 (CH_2), 67.98 (CH), 67.79 (CH), 67.03 (CH_2), 65.54 (CH), 50.63 (CH_2), 37.71 (CH_2), 21.12 (CH_3), 20.76 (CH_3), 20.67 (CH_3), 15.9 (CH_3). ^{19}F NMR (565 MHz, CDCl_3): δ -212.23 (t, $J = 47.7$ Hz).

3-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)-N-(3-(fluoromethyl)-4-(((2S,3S,4R,5S,6S)-3,4,5-trihydroxy-6-methyltetrahydro-2H-pyran-2-yl)oxy)phenyl)propanamide **AH062 (12)**



To a solution of compound **11** (158 mg, 0.246 mmol) in $\text{MeOH}/\text{CH}_2\text{Cl}_2$ (2/1) was added Na_2CO_3 (104 mg, 0.983 mmol) and the reaction mixture was stirred for 4h.

The reaction mixture was neutralised by H^+ resin, filtered and concentrated under reduced pressure. Purification by silica gel column chromatography (eluent: 99% to 94% $\text{PetEt}/\text{EtOAc}$ gradient + 1% AcOH) provided the desired final product (**AH062; 12**) (88.8 mg, 0.172 mmol) in 70% yield as a colourless solid. $R_f = 0.63$ ($\text{CHCl}_3/\text{MeOH} = 8/2$). ^1H NMR (600 MHz, CD_3OD): δ 7.62 (s, 1H), 7.53 (d, $J = 9.2$ Hz, 1H), 7.18 (d, $J = 8.9$ Hz, 1H), 5.56 (dd, $J = 87.2, 11.2$ Hz, 1H), 5.51 (d, $J = 3.4$ Hz, 1H), 5.48 (dd, $J = 87.2, 11.2$ Hz, 1H), 4.63 (s, NH), 4.06 (q, $J = 6.6$ Hz, 1H), 4.01 – 3.92 (m, 2H), 3.84 (t, $J = 6.0$ Hz, 2H), 3.76 (d, $J = 2.3$ Hz, 1H), 3.70 – 3.54 (m, 12H), 2.63 (t, $J = 6.0$ Hz, 2H), 1.20 (d, $J = 6.6$ Hz, 3H). ^{13}C NMR (151 MHz, CD_3OD): δ 170.85 (C), 132.76 (C), 121.54 (CH), 120.98 (CH), 114.92 (CH), 98.13 (CH), 79.85 (d, $J = 164.1$ Hz, CH_2F), 72.06 (CH), 70.23 (CH_2), 70.12 (CH), 70.09 (CH_2), 69.68 (CH_2), 68.27 (CH), 67.37 (CH), 66.80 (CH_2), 37.08 (CH_2), 15.22 (CH_3). ^{19}F NMR (565 MHz, CD_3OD): δ -215.71 (t, $J = 47.9$ Hz). HRMS (ESI-quadrupole) calculated for $\text{C}_{22}\text{H}_{33}\text{FN}_4\text{O}_9$ ($\text{M} + \text{H}$) $^+$: 517.2304; found: 517.2314.

Cloning, expression and purification of recombinant fucosidases BfFucH and AfcA

The codon-optimized genes encoding for BbAfcA (amino acid 577-1474, GenBank: AY303700) including the addition of GGSGGSHHHHHHH plus a stop codon and BfFucH (amino acid 21-434, GenBank: CAH08937.1) were synthesized by Genscript and subcloned into pMAL-c4x-1-H(RBS)+ (Aval/EcoRI) and pET47b+ (Acc65I/XhoI) respectively. BL21(DE3) (C2527H, NEB) cells were transformed with each of the individual vectors, a colony was picked from the 2xYT agar (BP97432, Fisher Bioreagents) plate with the appropriate antibiotic (ampicillin 100 μ g/mL, kanamycin 50 μ g/mL, Cayman chemicals) and expanded to a cell culture volume of 500 mL antibiotic (ampicillin 100 μ g/mL, kanamycin 50 μ g/mL) containing 2xYT medium at 37°C. The cells were induced at OD₆₀₀ = 0.6 with IPTG (final concentration was 1 mM, R0393, Thermo Scientific) and cultured overnight at 20°C. Then, cells were pelleted at 3000 xg, resuspended in 5% of culture volume with lysis buffer (1 mg/ml lysozyme (62971, sigma-aldrich) and 0.1% triton X-100 (T8787, Sigma-Aldrich), incubated at 37°C for 1 h and sonicated for 30 minutes on ice. Clear supernatant, by prior removal of cell debris by centrifugation at 10000 x g, was loaded onto a gravity flow 4 mL Ni-NTA column (17-5318-01, GE healthcare). A standard buffer consisting of 50 mM TRIS-HCl and 250 mM NaCl pH 8 was made. Imidazole was added to the standard buffer at concentrations of 20 mM, 50 mM or 250 mM to make wash 1, wash 2 and elution buffer (re-adjusted to pH 8 if needed) respectively. Once washed with 10 column volumes (CV) of wash 1, 10 CV of wash 2, the enzyme of interest was eluted in 3 CV of elution buffer. The eluate was concentrated using a Vivaspin 6 filter (VS0602, Sartorius) and further purified on a PBS equilibrated Superdex 200 Increase 10/300 GL column (28990944, GE healthcare) attached to a Shimadzu Nexera system using a flow of 0.45 ml/min collecting fractions each minute. Fractions containing pure enzyme were pooled, concentrated, aliquoted and stored at -20°C in 10% glycerol in PBS upon further use.

Chemoselective labelling

To study the limit of detection, various concentrations BfFucH (4 – 0.25 µg) or AfcA (8 – 3.7 µg) were incubated with 350 µM of **AH062**, then incubated for 1.5 h at 37 °C in PBS (pH 7.4). The negative controls included were a sample without any probe and a sample with the recombinant enzymes pre-incubated with the competitive inhibitor FNJ (final conc 100 µM, 99212-30-3, Carbosynth). The samples were mixed with 3x loading buffer, heated for 5 min at 95 °C and loaded onto a SDS-PAGE gel. After electrophoresis protocol of 10 min 90V, 50 min 200V the samples were blotted to a nitrocellulose membrane (TransBlot Turbo Transfer System, Biorad). The membrane was washed for 2 minutes in PBS after which it was incubated for 2h at room temperature in the presence of 100 µM alkyne-biotin (1262681-31-1, Jena Bioscience), 2.5 mM L-ascorbate (134-03-2, Sigma-Aldrich) and 0.5 mM Cu₂SO₄ (7758-99-8, Sigma-Aldrich) in PBS. The blots were washed 2 times 5 minutes with PBS containing 0.1% Tween-20 (PBS-T). The membrane was blocked using 5% skimmed milk in PBS-T with rocking for 1 h at room temperature. The blocking solution was decanted and a solution 1% skimmed milk in PBS-T containing α-biotin-HRP antibody conjugate (1:10.000, Jackson ImmunoResearch) was added and incubated for 1h at room temperature. Membranes were washed for 3 x 10 minutes with PBS-T. Detection of membrane bound αbiotin-HRP conjugates was accomplished by chemiluminescent detection using the Clarity Western ECL kit (Bio-Rad) and imaged in a Gel-Doc system (Bio-Rad). A second gel with the same samples was run in parallel for each experiment, to which silver staining or PageBlue staining was applied to provide loading control according to the protocol by Chevallet et al.²⁸

Labelling selectivity assay

BfFucH (960 nM) was incubated with 350 µM **AH062** in the presence of decreasing concentrations of TF (960 - 60 nM) for 2 h at room temperature in PBS. The enzyme mixtures were transferred, conjugated to alkyne-biotin, and visualized as described above.

Mass spectrometry

For the peptide centric LC-MS-MS, 5 µg expressed fucosidases either labelled or unlabelled were denatured and reduced using 5 mM tris(2-carboxyethyl)phosphine (TCEP) and incubated for 15 min. at 60 °C. Samples were subsequently alkylated with 20 mM chloroacetamide (CAA) and digested proteolytically using pepsin (Roche) in a 0.2% trifluoroacetic acid (TFA) buffer at pH <2 (1:75 enzyme:protein ratio) and incubated overnight at 37 °C. The protein was subjected to several proteases including trypsin, chymotrypsin and glu-C (data not shown). Pepsin proved to be the most suitable for producing peptides at appropriate lengths with the highest sequence coverage. Following digestion, solid phase extraction was performed using Oasis microElution 96-well plates (Waters, Wexford, Ireland). In short, wells were conditioned with acetonitrile (ACN), then equilibrated with 0.1% TFA. Supernatant from the digestion was loaded and washed twice with 0.1% TFA. Peptides were eluted using 60% ACN 0.1 % TFA. Peptide

eluates were dried by low-pressure rotary evaporation and resuspended in 2% formic acid for LC-MS² analysis.

Per peptide sample, 50 ng was analyzed using a Thermo Scientific Ultimate HPLC nanoflow system, hyphenated to an Orbitrap Exploris 480 Mass Spectrometer (Thermo Fisher Scientific). Buffer A (0.1% formic acid (FA) in water) and buffer B (0.1% FA, 20% ACN in water) were used for peptide chromatography with the following gradient: 0 min. 9% B; 2 min. 13% B; 39 min. 44% B; 42 min. 99% B; 46 min. 99% B; 47-55 min. 9% B. Mass spectrometry analysis was performed in positive ion mode using electrospray ionization from a coated fused silica emitter at 1900 V spray voltage. For MS¹ scans, the mass range was set from m/z 375 to 1600 with a resolution of 60,000. The AGC target was set to the MS standard with automated maximum injection times. The data dependant MS² acquisition method initiated HCD fragmentation at 28% of normalized collision energy on the highest charge state and lowest m/z signals within a 1 s cycle time. The exclusion time was set to 10 s. Resolution for MS² acquisition was at 15,000 with a mass range from m/z 40 to 2500. Here, the AGC target was also set to standard with automated maximum injection times. For data analysis of LC-MS² we used the sequence of recombinantly expressed fucosidase from *T. maritima* and *B. fragilis*. Data was interpreted with Byonic v3.11.3 (Protein Metrics Inc.). Raw data were searched non-specifically with regards to digestion specificity. Precursor mass tolerance was set to 10 ppm and HCD fragment mass tolerance to 20 ppm. As a fixed modification, Cys carbamidomethylation was included in the search. Met and Try oxidation was added as a common variable modification. The possible masses of the probe, protonated (351.17 Da) and deprotonated (350.16 Da) were searched against the known sequences of the enzyme (See Figure S4. protein sequences Bfuch and TmFuc).

For protein-centric LC-MS analysis, samples containing labelled and unlabelled fucosidase from *T. maritima* and *B. fragilis* were chromatographically separated on a Thermo Scientific Vanquish Flex UHPLC system. 8 μ g of protein material was loaded on a reverse phase column heated to 80 °C. LC-MS runtime was set to 22 min with a flow rate of 250 μ L/min. Analytes were eluted using two mobile phases, 0.1% FA in water (buffer A) and 0.1% FA in ACN (buffer B) with a gradient of 25%-46% B for 14 minutes. The LC-MS analysis was performed on an Orbitrap Exploris 480 Mass Spectrometer expanded with intact (biopharma) mode. Data was collected with the instrument set to intact protein mode, at a resolution of 7,500 at m/z 200. Raw spectra were charge state deconvoluted using PMI intact mass software at a m/z range of 500-2000, using charge vector spacing of 0.6 with a baseline radius of a m/z of 15. Mass smoothing sigma was set to 3 and mass spacing to 0.5 and a maximum of 10 iterations. Annotations by the program were checked manually using the raw data.

Fluorescent microscopy and analysis

Bacteroides fragilis (NTCT9343) was grown anaerobically at 37°C in GAM medium (Gifu Anaerobic Broth, 05422, HyServe). Overnight cultures (18h) were collected via centrifugation (4000 rpm, 5 min, 4 °C) resuspended in PBS and 500 µL of OD₆₀₀ 1.0 was transferred into aliquots. The aliquots were centrifuged (4000 rpm, 5 min, 4 °C) and resuspended in 100 µL PBS containing the competitive inhibitor FNJ (final = 100 µM) or an equal volume of PBS. Bacteria were incubated for 30 minutes at 37°C. The appropriate samples were treated with 350 µM AH062 and incubated for 1.5h at 37°C. Bacteria were collected via centrifugation at 8000 rpm for 5 minutes and washed 2x with 1 mL PBS. Pellets was resuspended in 100 µL CuAAC reaction buffer (1 µM alkyne-Alexa488 (Jena Bioscience), 0.5 mM CuSO₄, 2.5 mM L-ascorbate in PBS). Bacteria were incubated in the dark at rt for 3 hours and collected by centrifugation as above. Bacteria were washed 3x by resuspension in PBS-T, incubating for 3 minutes in the dark, and centrifugation as above. The bacteria were resuspended in 50 uL PBS and labelled by 1 µL CTY (Cell trace yellow cell proliferation kit, Invitrogen) per vial for 20 minutes at 42 °C. The bacteria were washed 2x times with PBS followed by fixation with 4% PFA 20 minutes rt. The bacteria were washed an additional 2 times in PBS. The bacterial pellets were resuspended in Prolonged Diamond solution (Thermo Fisher) and mounted onto cover slips for confocal imaging.

Images were either collected on a Leica SPE-II confocal microscope using a 100x objective (HCX PL APO CS 100.0x1.40 oil) controlled by Leica LAS AF software with default settings to detect Alexa488 and CTY.

Or images were collected on an OMX-V4 BLAZE super-resolution microscope using a 60x SIM lens (Olympus U-PLAN APO, NA 1.42, WD=0.15). The raw SIM data were reconstructed in SoftWoRx (Cytiva) and channel alignment was performed using Chromagnon using reference images of calibration tetraspec beads (Thermo Scientific).²⁹ Images were converted and deconvolved (blind, iteration 20x) in NIS elements (NIKON). Maximum intensity projections are shown.

Flow cytometry

Bacteria were prepared similar to the procedure described for “Fluorescent spectroscopy and analysis” until the CTY labelling step. Pellets of clicked bacteria were fixed by the addition of 500 µL 4% PFA and incubated for 20 minutes at rt. Bacteria were washed 2x with PBS and taken up in 1 mL aquadest. Bacteria was diluted an additional 10x in aquadest and analysed using a BD FACSVerser. Flow cytometry data was analysed using FlowJo 10.

Acknowledgements

We thank Richard Wubbolts and Esther van 't Veld of the Centre for Cell Imaging (CCI) of the Faculty of Veterinary Medicine for technical support, Gerlof Bosman for the expression of the recombinant fucosidases, and Dario Cramer and Karli Reidinger for the MS analysis and interpretation.

References

- 1 K. Wang, W. Guo, N. Li, J. Shi, C. Zhang, W. Y. Lau, M. Wu and S. Cheng, *Br. J. Cancer*, 2014, **110**, 1811–1819.
- 2 T. C. Cheng, S. H. Tu, L. C. Chen, M. Y. Chen, W. Y. Chen, Y. K. Lin, C. T. Ho, S. Y. Lin, C. H. Wu and Y. S. Ho, *Oncotarget*, 2015, **6**, 21283–21300.
- 3 T. W. Liu, C. W. Ho, H. H. Huang, S. M. Chang, S. D. Popat, Y. T. Wang, M. S. Wu, Y. J. Chen and C. H. Lin, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 14581–14586.
- 4 M. Nandakumar, Y. L. Hsu, J. C. Y. Lin, C. Lo, L. C. Lo and C. H. Lin, *ChemBioChem*, 2015, **16**, 1555–1559.
- 5 H. Wu, O. Rebello, E. H. Crost, C. D. Owen, S. Walpole, C. Bennati-Granier, D. Ndeh, S. Monaco, T. Hicks, A. Colville, P. A. Urbanowicz, M. A. Walsh, J. Angulo, D. I. R. Spencer and N. Juge, *Cell. Mol. Life Sci.*, 2020, **1**, 3.
- 6 Y. Luijckx, N. Bleumink, J. Jiang, H. S. Overkleeft, M. Wösten, K. Strijbis and T. Wennekes, *Cell. Microbiol.*, 2020, cmi.13252.
- 7 J. M. Garber, H. Nothhaft, B. Pluvinaige, M. Stahl, X. Bian, S. Porfirio, A. Enriquez, J. Butcher, H. Huang, J. Glushka, E. Line, J. A. Gerlt, P. Azadi, A. Stintzi, A. B. Boraston and C. M. Szymanski, *Commun. Biol.*, 2020, **3**, 1–11.
- 8 J. M. Pickard and A. V Chervonsky, *J. Immunol.*, 2015, **194**, 5588–5593.
- 9 M. T. Sorbara and E. G. Pamer, *Mucosal Immunol.*, 2019, **12**, 1.
- 10 M. Stahl, L. M. Friis, H. Nothhaft, X. Liu, J. Li, C. M. Szymanski and A. Stintzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 7194–7199.
- 11 T. Katayama, A. Sakuma, T. Kimura, Y. Makimura, J. Hiratake, K. Sakata, T. Yamanoi, H. Kumagai and K. Yamamoto, *J. Bacteriol.*, 2004, **186**, 4885–4893.
- 12 J. You, S. Lin and T. Jiang, *Front. Microbiol.*, 2019, **10**, 1756.
- 13 Y. Xu, N. Uddin and G. K. Wagner, *Covalent Probes for Carbohydrate-Active Enzymes: From Glycosidases to Glycosyltransferases*, Elsevier Inc., 1st edn., 2018, vol. 598.
- 14 L. Wu, Z. Armstrong, S. P. Schröder, C. de Boer, M. Artola, J. M. Aerts, H. S. Overkleeft and G. J. Davies, *Curr. Opin. Chem. Biol.*, 2019, **53**, 25–36.
- 15 S. G. Withers, K. Rupitz and I. P. Street, *J. Biol. Chem.*, 1988, **263**, 7929–7932.
- 16 J. Jiang, W. W. Kallemijn, D. W. Wright, A. M. C. H. C. H. Van Den Nieuwendijk, V. C. Rohde, E. C. Folch, H. Van Den Elst, B. I. Florea, S. Scheij, W. E. Donker-Koopman, M. Verhoek, N. Li, M. Schürmann, D. Mink, R. G. Boot, J. D. C. C. Codée, G. A. Van Der Marel, G. J. Davies, J. M. F. G. F. G. Aerts, H. S. Overkleeft, V. Coco Rohde, E. Colomina Folch, H. Van Den Elst, B. I. Florea, S. Scheij, W. E. Donker-Koopman, M. Verhoek, N. Li, M. Schürmann, D. Mink, R. G. Boot, J. D. C. C. Codée, G. A. van den Marel, G. J. Davies, J. M. F. G. F. G. Aerts and H. S. Overkleeft, *Chem. Sci.*, 2015, **6**, 2782–2789.
- 17 M. Kurogochi, S. I. Nishimura and Y. C. Lee, *J. Biol. Chem.*, 2004, **279**, 44704–44712.
- 18 L. M. Chauvigné-Hines, L. N. Anderson, H. M. Weaver, J. N. Brown, P. K. Koech, C. D. Nicora, B. A. Hofstad, R. D. Smith, M. J. Wilkins, S. J. Callister and A. T. Wright, *J. Am. Chem. Soc.*, 2012, **134**, 20521–20532.
- 19 Y. Hsu, M. Nandakumar, H. Lai, T. Chou, C. Chu, C. Lin and L. Lo, *J. Org. Chem.*, 2015, **80**, 8458–8463.
- 20 H. Hinou, M. Kurogochi, H. Shimizu and S. I. Nishimura, *Biochemistry*, 2005, **44**, 11669–11675.
- 21 K. Noguchi, T. Shimomura, Y. Ohuchi, M. Ishiyama, M. Shiga, T. Mori, Y. Katayama and Y. Ueno, *Bioconjug. Chem.*, 2020, **31**, 1740–1744.

- 22 D. H. Kwan, H.-M. Chen, K. Ratananikom, S. M. Hancock, Y. Watanabe, P. T. Kongsaree, A. L. Samuels and S. G. Withers, *Angew. Chemie*, 2011, **123**, 314–317.
- 23 S. H. L. Verhelst, K. M. Bongers and L. I. Willems, *Molecules*, 2020, **25**, 5994.
- 24 G. Baisch and R. Öhrlein, *Bioorg. Med. Chem.*, 1997, **5**, 383–391.
- 25 E. D. Goddard-borger and R. V Stick, *Org. Lett.*, 2007, **9**, 3797–3800.
- 26 G. M. Loudon and D. E. Koshland, *J. Biol. Chem.*, 1970, **245**, 2247–2254.
- 27 T. I. Tsai, S. T. Li, C. P. Liu, K. Y. Chen, S. S. Shivatare, C. W. Lin, S. F. Liao, C. W. Lin, T. L. Hsu, Y. T. Wu, M. H. Tsai, M. Y. Lai, N. H. Lin, C. Y. Wu and C. H. Wong, *ACS Chem. Biol.*, 2017, **12**, 63–72.
- 28 M. Chevallet, S. Luche and T. Rabilloud, *Nat. Protoc.*, 2006, **1**, 1852–1858.
- 29 A. Matsuda, L. Schermelleh, Y. Hirano, T. Haraguchi and Y. Hiraoka, *Sci. Rep.*, 2018, **8**, 7583.

Supplementary Figures

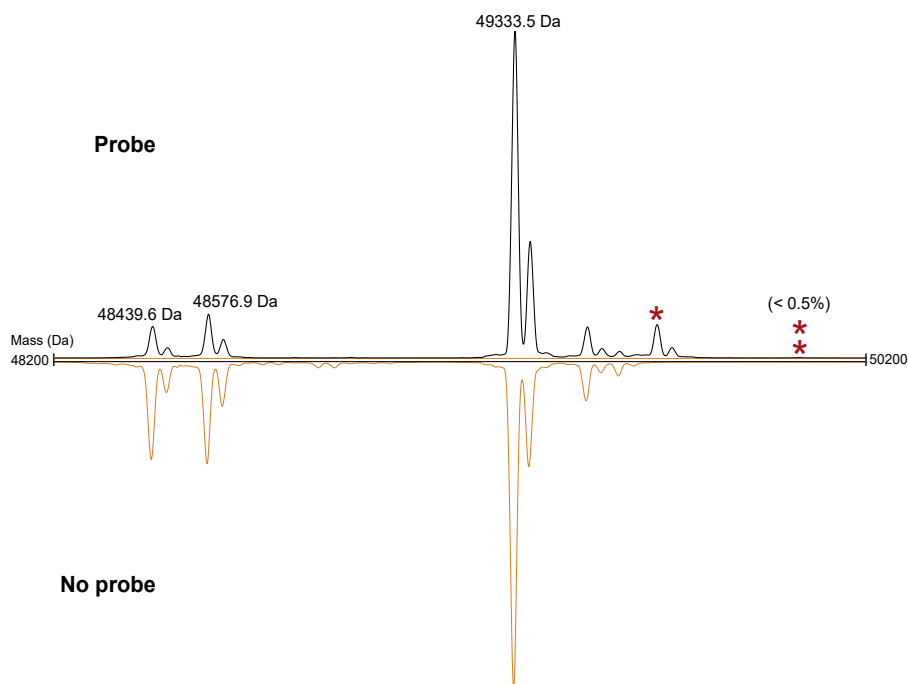
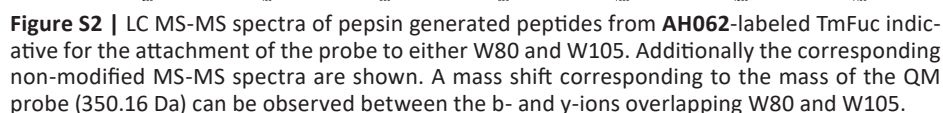


Figure S1 | Zero-charge deconvoluted mass spectra of the intact BfFucH analyzed by LC-MS. Shown here are the fucosidase without labelling (top) and after labelling with the **AH062** probe (top). Three variants of the BfFucH were observed, one at a retention time of 9.2-9.6 minutes, representative of the intact molecule (49333.5 Da) and two at a retention time of 9.6-10.1 minutes, representative of a truncated molecule (-756 Da; -893 Da). Red asterisks indicate labeling with the **AH062** probe, corresponding to an approximate mass increase of 350 Da or 700 Da (doubly labeled).



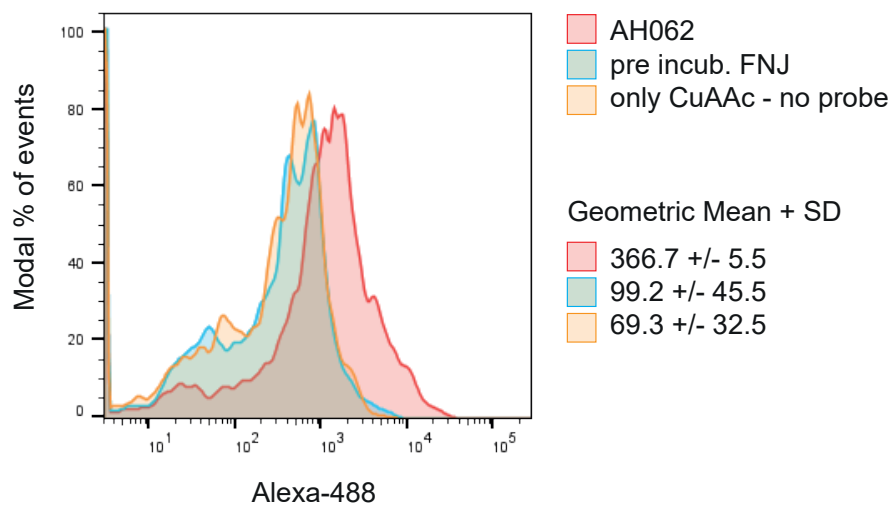


Figure S3 | Flow cytometry analysis of *B. fragilis* labelled with **AH062** and tagged with Alexa-488. Image shows representative of the experiment. Geometric mean and SD is reported for each population, n=3.

BfFucH

10 20 30 40 50 60
 QKQYQPT^WEAN LKARSEFQDN KFGIFLH^WGL YAMLATGE^WT MTNNNLNYKE YAKLAGGFYP
 70 80 90 100 110 120
 SKFDADK^WVVA AIKASGAKYI CFTTRHHEGF SMFDTKYSYD NIVKATPFKR DVVKELADAC
 130 140 150 160 170 180
 AKHGIKLHFY YSHID^WYRED APQGR^TGRRT GRPNPKGDMA HHHHHHSAAL EVLFQ^WGKSY
 190 200 210 220 230 240
 YQFMNNQLTE LLTNYGPIGA I^WFDG^WWDQD INPDFD^WELP EQYALIHRLQ PACLVGN^WNHH
 250 260 270 280 290 300
 QTPFAGEDIQ IFERDLPGEN TAGLSGQSVS HLPLETCE^TMG NGM^WG^WYK^WITD QNYKSTK^TLI
 310 320 330 340 350 360
 HYLVAAGKD ANLLMNIGPQ PDGELPEVAV QRLKEVGE^WWM SKYGETIYGT RGLLVAPHD^W
 370 380 390 400 410 420
 GVTQKGNKL YVHILNLQDK ALFLPIVDKK VKKAVVFADK TPVRFTKNKE GIVLELAKVP
 430
 TDVDYVVELT ID

TmFuc

10 20 30 40 50 60
 MISMKPRYPK D^WWESLREHTV PK^WWFDKAKFG IFIH^WG^WGIYSV PG^WWATPTGEL GKVPMDA^WWFF
 70 80 90 100 110 120
 QNPYA^WEYEN SLRIKESPT^WY EYHVKT^WYGEN FEYEKFADLF TAEK^WWDPQ^WE^WW ADLFKKAGAK
 130 140 150 160 170 180
 YVIPTTKHHD GFCL^WG^WTKYT DFNSVKRGPK RDLVGDLAKA VREAGLRFGV YYSGLD^WWRF
 190 200 210 220 230 240
 TTEPIRYPED LSYIRPNTYE YADYAYKQVM ELVDLYLPDV L^WWNDMG^WWPEK GKEDLKYLFA
 250 260 270 280 290 300
 YYYNKHPEGS VNDR^WWGVPH^W DFKTAEYHVN YPGDLPGYK^WW EFTRGIGLSF GYNRNEGPEH
 310 320 330 340 350 360
 MLSVEQLVYT LVDVVS^WKGGN LLLNVGPKGD GTIPDLQKER LLGLGE^WWLRK YGDAIYGT^WSV
 370 380 390 400 410 420

Figure S4 | Protein sequences BfFucH and TmFuc. W residue in bold.

Chapter 6

**Summary, Future Perspectives and
Concluding Remarks**

6.1 Summary

As for all cells on earth, both human gut epithelial cells and our gut microbiota are covered by a dense coat of glycans, called the glycocalyx. The unique glycans in this glycocalyx are not encoded in the genome and the biosynthesis of their complex structures is not template-driven. The regular molecular biology tools to study and manipulate the biomolecules DNA and proteins at the molecular level can therefore not be easily applied to elucidate the functional role of glycans and their interacting proteins at this human-gut microbiota interface. The application of techniques from the field of chemical biology however has resulted in a successful strategy for this through the development of smart tailor-made probes – carbohydrate-based bioactive small molecular tools – that target a specific carbohydrate or interacting protein class in a cell or whole organism. In the past decades such glycan-based probes have increasingly been used to unravel glycan-related biological processes. The focus of the research reported in this thesis was on studying fucosylated glycans and their interacting enzymes at the human-gut microbiota interface. Fucosylated glycoproteins are abundant at the gut-microbiota interface and have been implicated in critical biological processes such as immune response, signal transduction, and adhesion of pathogens. Quantification and visualization of fucosidases, the enzymes involved in altering this fucosylation pattern, will help us unravel their biological importance. The research described in this thesis includes the development of two different fucosidase-targeting activity-based probes (ABPs). These new tools can in the future potentially be used in high throughput methods to study the relationship between fucosylation and the maintenance of homeostasis at the human-gut microbiota interface.

In **Chapter 1**, a concise and up to date overview of scientific literature on the role of fucose and fucosidase enzymes at the human-gut microbiota interface is provided and discussed. In addition, the chapter provides a brief overview of the mechanism behind the established ABPs and highlights their current use. In **Chapter 2**, we reported on cross-feeding mechanisms underlying *Campylobacter jejuni* 108 growth and invasion. Although the *C. jejuni* 108 strain lacks fucosidase activity, its genome does show the presence of a L-fucose operon. *C. jejuni* 108 showed increased growth when grown in the presence of the fucosidase expressing gut commensal *Bacteroides fragilis* combined with Porcine Gastric Mucin (PGM). A mucin-expressing cell-based model in combination with a known fucosidase inhibitor was used to study the effect of *B. fragilis* fucosidases on the invasion of *C. jejuni* 108. *C. jejuni* 108 showed to be more invasive in the presence of *B. fragilis* and this increase was due to fucosidase activity. Therefore, it was concluded that *C. jejuni* 108 strain is dependent on exogenous fucosidases for increased growth and invasion.

Butyrate is a compound that is continuously produced by the microbiota that mediates human gut microbiota homeostasis, regulates glycosylation related genes, and induces differentiation. The Caco-2 intestinal cell line model is widely used to study bacterial interactions with specific *O*-glycans, but no studies had yet investigated the effect of the bacterial metabolite butyrate on cellular differentiation and its *O*-glycan composition. In **Chapter 3**, we applied a PGC nano-LC–ESI-MS/MS method to enable identification and quantification of *O*-glycan alterations in differentiating Caco-2 cells over a 24-day period in response to butyrate stimulation. Several specific glycans, such as the sDa antigen and di-sialylated structures showed to be higher expressed in the presence of butyrate in comparison to spontaneously differentiated Caco-2 cells in the absence of butyrate. Additionally, later time points in spontaneous differentiation showed a stabilization of the glycome. In contrast, butyrate stimulation did not result in stabilization of the glycome. In these later time points, several glycans carrying blood group antigens terminated with α 1,2-fucose showed distinct alterations in the butyrate-stimulated groups. This study provides new insights in the *O*-glycan composition of intestinal cells that are differentiated in the presence of the naturally abundant bacterial compound butyrate

In the following chapters, we focussed our attention on the role of L-fucose at the human-gut microbiota interface. An increasing amount of research suggests a role for bacterial fucosidases in either maintaining or tipping the balance in gut homeostasis. Therefore, **Chapter 4** focussed on expanding the toolbox of ABPs capable of targeting GH29 fucosidases. We established a facile synthesis route for a novel 2-deoxy-2-fluoro based fucosidase probe and determined the effect of the C6-azide on its inhibition properties by a kinetic study compared to the known 2FFucF inhibitor. In addition, this new ABP was shown to be capable of targeting bacterial GH29 fucosidases with different linkage specificity.

In the introduction of Chapter 1 we highlighted the fact that there are currently two classes of fucosidases, of which only the retaining fucosidases are currently targeted by ABPs. In **Chapter 5**, we describe the synthesis of an *ortho*-quinone methide based probe inspired by a compound previously published by Hsu et al.¹ We show that besides being able to target retaining GH29 fucosidases this probe enables visualization of inverting GH95 fucosidase activity. We established that this probe is compatible with high throughput methods such as flow cytometry. In addition, top-down and bottom-up MS approaches allowed for unique insights into possible sites of attachment of the activated quinone methide functionality.

6.2 Future perspectives on fucosidase ABPs

The human intestinal epithelium is in close contact with a vast number of diverse bacteria and at this complex gut-microbiota interface fucosidases play an essential role in the maintenance of homeostasis. For a more detailed discussion on the role of fucosidases in homeostasis see **Chapter 1**. These crucial fucosidases are widely abundant, show a broad range of specificities, but due to the microbiota complexity are as of yet understudied with respect to their identities, spatio-temporal activity and functional roles. Thus far, there are only a limited number of ABPs known to be capable of targeting the bacterial fucosidases. The research described in this thesis added two additional probes to the library. However, optimization and development of new fucosidase-targeting ABPs is necessary to decipher the different functions of L-fucosidases at the complex human-gut microbiota interface. The next two sections discuss two interesting directions towards achieving the goal of further expanding the molecular toolbox for studying fucosidases.

Tuning the leaving group of 2-deoxy-2-fluoro glycoside based fucosidase inhibitors

One of the challenges in the development of activity-based probes based on 2-deoxy-2-fluoro glycosides is their relatively slow rate of formation of the covalent inhibitor-enzyme intermediate. The design of the inhibitor is based on the ability to stabilize the inhibitor-enzyme intermediate by introduction of an electron-withdrawing fluorine at the C-2 position that slows down or prevents the hydrolysis of intermediate's ester linkage.² However, the electron withdrawing fluorine on the C-2 position also reduces the formation rate of the inhibitor-enzyme intermediate. Increasing the enzyme-inhibitor intermediate rate of formation can be achieved by tuning the leaving group on the anomeric position of the 2-deoxy-2-fluoro glycoside. The 2-deoxy-2-fluoroglycoside **2FFuc** with a fluorine leaving group, was evaluated as ABP by us (**Chapter 4**, Figure 1a). Introduction of an azide at the C-6 position gave the two-step labelling probe **YL209**. This azide functionalized 2-deoxy-2-fluoro glycoside labelled >70% of the fucosidases BfFucH, provided that prolonged reaction times (>4h) and relatively high concentrations (50 mM) were used. Increasing the formation rate of the inhibitor-enzyme intermediate can be achieved by tuning the leaving group at the anomeric position of the 2-deoxy-2-fluoroglycoside probes. This was previously studied by Walvoort *et al.* who synthesized and evaluated different leaving groups on the anomeric position of their GBA-1 inhibitor.³ It was shown that the probe with either a phosphate or an imidate leaving group labelled GBA-1 more efficiently than the fluoride probe (Figure 1b). Furthermore, they showed the imidate probe to be more hydrolytically stable than the phosphate probe.³ The finding that the inhibitory potency of the 2-deoxy-2-fluoro glycoside could be fine-tuned by the installation of an imidate group on the anomeric position was an incentive to further explore this *N*-phenyl trifluoroacetamide imidate functionality for our previously described 2-deoxy-2-fluoro based α -L-fucosidases inhibitor and ABP (Figure 1c).

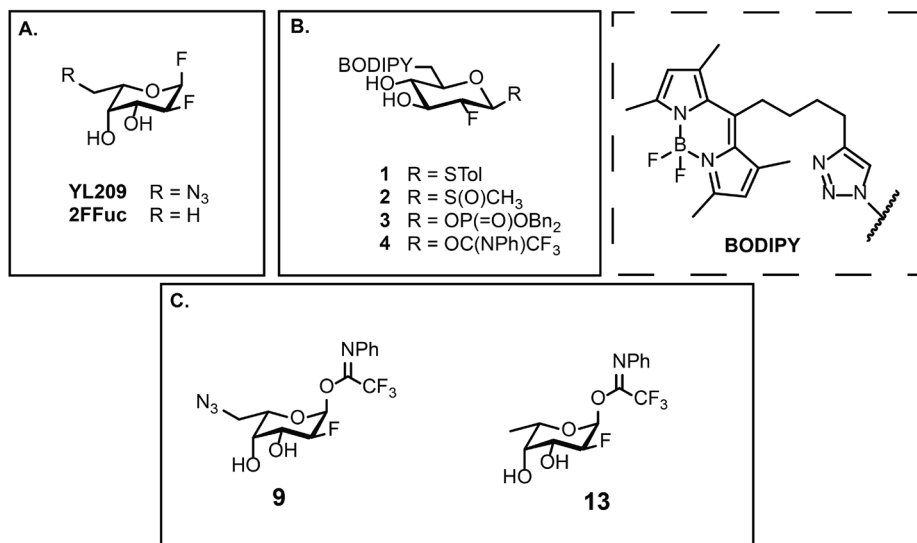
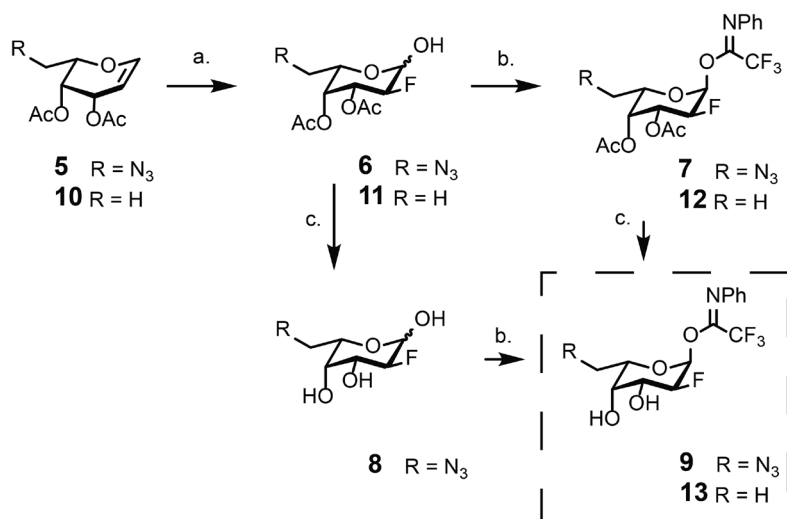


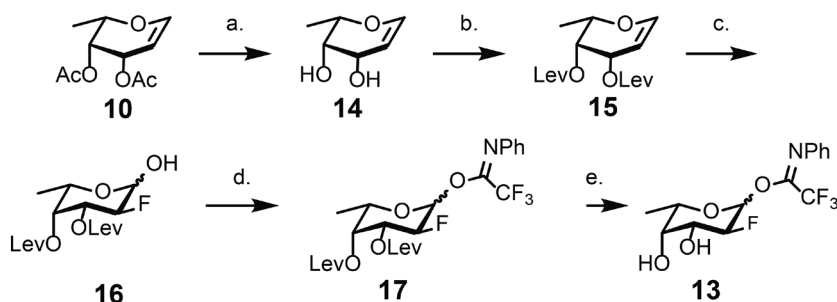
Figure 1 | a) Modified 2-deoxy-2-fluoro glucosyl fucosidases ABP **YL209**. b) 2-Deoxy-2-fluoro glucosyl probes with varying leaving groups according to Walvoort *et al.* c) 2-Deoxy-2-fluoro fucosyl ABP and inhibitor, respectively **9** and **13**.

Both *N*-phenyl trifluoroacetimide imidate inhibitors (**9** and **13**) were accessed through a similar synthesis route that is depicted in Scheme 1. Peracetylated glucal **5** and **10** were used as the starting compounds. The glucals were treated with Selectfluor® to provide only the correct configuration, the formation of the talo-epimer was not observed. The imidates were introduced on the 2-fluoro glycosides using *N*-phenyl trifluoroacetimidoyl chloride in combination with K₂CO₃. In the final step **7** and **12** were deacylated under Zémlen conditions. Unfortunately, in our hands we were not able to isolate the fully deprotected 2-fluoro-glycoside imidates (**9** and **13**). Parallel to this approach it was also attempted to selectively install the imidate leaving group on the fully deprotected 6-azido-2-fluoro glycosides (**8**). TLC analysis showed the formation of several products. Unfortunately, in our hands we were not able to detect the fully deprotected 6-azido-2-fluoro-glycoside imidate (**9**) among the possible fractions obtained after performing a silica column.



Scheme 1 | Synthesis of 2-deoxy-2-fluoro fucosyl ABP and inhibitor, **9** and **13**. *Reagents and conditions:* (a) Selectfluor®, acetone/H₂O (7:1) **6**: 73%, **11**: 53%; (b) K₂CO₃, ClC(=NPh)CF₃, acetone **7**: 63%, **12**: 87%; (c) NaOMe, MeOH, no detectable traces of **9** or **13**, quant. yield of **8**.

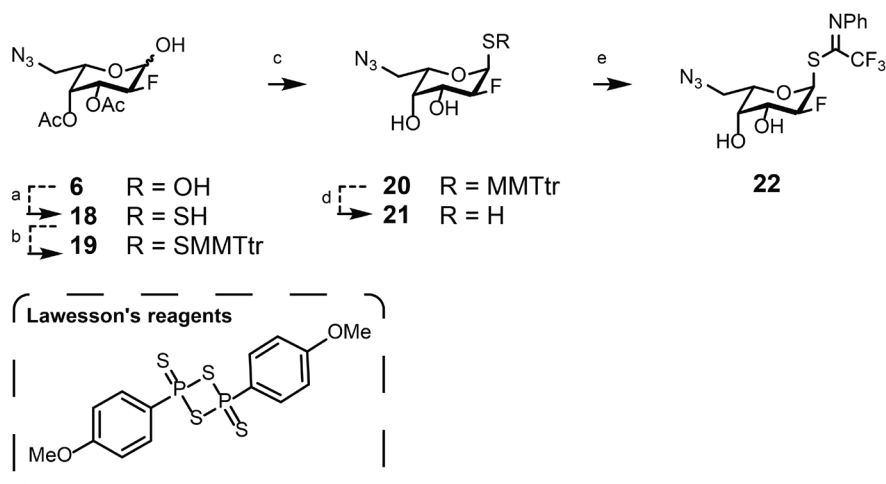
To solve this issue, we evaluated a different protection group strategy with milder deprotection conditions. The levulinate esters are widely used in oligosaccharide synthesis, due to their mild and selective deprotection by hydrazine acetate. The *N*-phenyl trifluoroacetimidate imidate inhibitor was accessed through a similar route starting from the peracetylated L-fucose (Scheme 2). Deacetylation using NaOMe in MeOH gave 2-fluoroglycoside **14**. Esterification of the C-3 and C-4 hydroxyl groups in **14** with levulinic acid provided **15** in 76% yield. Intermediate **15** was treated with Selectfluor® to provide **16** as an anomeric mixture. Unfortunately, the desired product could not be purified by column chromatography. We therefore continued with the α/β mixture. The anomeric imidate was introduced on the 2-fluoro glycosides using the *N*-phenyl trifluoroacetimidoyl chloride in combination with K₂CO₃ to provide **17**. An excess of hydrazine acetate was used to remove the levulinoyl groups, which yielded multiple products. Unfortunately, TLC-MS analysis did not indicate the formation of the desired product (**12**). This result might be explained by the relatively high instability/reactivity of the fucosyl and galactosyl imidates as was previously reported by Jensen and Bols.⁴ It was postulated that the effect imposed by an electron withdrawing group at C-4 is less when placed axially than equatorially.



Scheme 2 | Synthesis of 2-deoxy-2-fluoro fucosyl inhibitor **12** by applying a different protecting group strategy. *Reagents and conditions:* (a) NaOMe, MeOH **14**: quant. yield; (b) levulinic acid, EDC, DMAP, dry DCM **15**: 76%; (c) Selectfluor®, acetone/H₂O (7:1) **16**: 49%; (d) K₂CO₃, ClC(=NPh)CF₃, acetone **17**: 68% (e) hydrazine acetate, DCM/MeOH (1:1) no detectable traces of **12**.

One alternative that might still work is the use of a thioimide as a more stable leaving group as they can withstand many reaction conditions associated with protecting group manipulations.^{5,6} Scheme 3 depicts a possible route for the synthesis of a 2-deoxy-2-fluoro-2-galactosyl thioimides. The 2,6-dideoxy-2-fluoro-6-azido fucose derivative **6** is converted to the corresponding thiol **18** with the help of Lawesson's reagent.⁷ Protection of the thiol function with an monomethoxytrityl (MMTtr) group should allow for separation of α -**19** and β -**19**.⁸ Deacetylation and deprotection of the MMTtr group is followed with introduction of the *N*-phenyl trifluoroacetimide imide under basic conditions to give the potential thioimide probe **22**.⁵

The proposed 2-deoxy-2-fluoro-galactosyl thioimide ABP may have an improved inhibitory profile thereby potentially expanding the possibilities of this probe for *in vivo* studies. Another potential future application of these ABPs may be found in the development of potent chemical chaperones to stabilize the 'native' fold(ing). Fucosidosis, a lysosomal storage disease caused by mutations in the FUCA1 gene, is an example of a disease that might benefit from the treatment with potential chaperones.^{9,10} Semi-irreversible ABPs such as 2-deoxy-2-fluoro-glycosides, equipped with optimized leaving groups as discussed previously are an attractive candidate for this. These types of ABPs should only transiently bind to the active site catalytic nucleophile, thereby stabilising the catalytically active correctly folded enzyme conformation, promoting the critical process of folding. Afterwards the rescued enzyme will reactivate following the slow irreversible release of 2-deoxy-2-fluoro-L-fucopyranose.



Scheme 3 | Proposed synthesis of thioimide fucosidases probes. *Reagents and conditions:* (a) Dioxane, Lawesson's reagent, 80 °C; (b) MMTtrCl, pyridine; (c) NaOMe/MeOH; (d) TFA, TES-H, DCM; (e) ClC(=NPh)CF₃, K₂CO₃, acetone.

Exploiting the hydrophobic concavity in fucosidases near the C6 methyl group of L-fucosides – new ABPs

Structure determinations and fucosidase-inhibitor complexes have yielded considerable insights into the topography of the active site and mechanism of retaining fucosidases.^{11,12} Several three-dimensional structures of GH29 α -L-fucosidases solved by X-ray diffraction of crystals are available in the Protein Databank (PDB). Among which the inhibitor bound structure of *Thermotoga maritima* GH29 α -L-fucosidase TM0306 (TmFuc), the inhibitor bound structure of *Bacteroides thetaiotaomicron* GH29 α -L-fucosidase BT2970 (BtFuc2970), and the structures of *Bifidobacterium longum subsp. infantis* GH29 α -L-fucosidase BiAfcB Blon2336 (BIFuc) and *Bacteroides thetaiotaomicron* GH29 α -L-fucosidase BT2192 (BtFuc2192).

Recent research suggests that the GH29 family can be divided into two subfamilies (A and B) based on phylogenetic relationships and substrate specificities. Subfamily A consist of fucosidases that have broad substrate specificity, whereas subfamily B fucosidases are more specific to α -(1,3) or α -(1,4) linkages.¹³ Figure 2 provides an alignment of representative sequences from both subfamilies, highlighting some important features.

GH29B	BinfantisFuc	VRPSSRLAWORMEMYAFLHFGMNTMDREWGLG-----	50
	BtFuc1625	LVPTPQAAWQOMELTAFLHFGINTFTGREWGDG-----	82
	BtFuc4136	LVPTPQVVWQOMELTAFLHFSINAFTHGREWGDG-----	82
	BtFuc2970	KRQDPAMQKERNRLGAFIHWGLYAIPEGEMNGKVYGG-----AAEWLKSW--	92
GH29A	TmFuc	LREHTVPKKEKAKFGFIHWGIYSVPGWATPTGELGKVPMDAWFFQNPYAENWYENSLEI	74
	Hs-FucA1	LDSRPLPAWDEAKFGVFIHWGVFSVPAWGSE-----WFWWH-----QG	78
		* *	
GH29B	BinfantisFuc	HEDE-----ALFNERNVDVDMMDALVAGGMAGVILTKKHHDGFC	90
	BtFuc1625	QEDF-----AIFNPTELDAEQWVRTLKDAGFKMVLLTAKHHHDGFC	122
	BtFuc4136	KEDF-----NLFNPTELDAEQWVSNLKEAGFKMVLLTAKHHHDGFC	122
	BtFuc2970	-----AKVPADEWLKLMQDWNPTKFDKAKWAKMAKEMGTKYVKIITKKHHEGFC	140
GH29A	TmFuc	KESPT-WEYHVKTGYGENFYEYKFDLFTAEKWDPOENADLFKKAGAKYVITPTKKHHDGFC	133
	Hs-FucA1	EGREYQYQRFMRDNYPPGFSYADFGPQETARFFHPPEENADLQAAAGAKYVILTLKKHHEGFT	138
		* * * * *	
GH29B	BinfantisFuc	LWPSRL-TRHTVASSPWRECKGDLVREVSERARRHGLKFGVYLS---PWRDTEESYK--	144
	BtFuc1625	LWPTAT-TKHSVASSPWKNQGDVVKELRNACDKYDMKFGVYLS---PWRDRAECYGD--	176
	BtFuc4136	LWPTAT-TKHSVASSWKNQGDVVKELRKACKKYGMRFGVYLS---PWRDRAECYGD--	176
	BtFuc2970	LWPSKY-TKYTVANTP---YKRDILGELVKANDEGIDVHFYFS-VMDNSNPDYRDIKS	195
GH29A	TmFuc	LWGTYK-TDFNSVKRG---PKRDLVGLDAKAVREAGLRFGVYSGGLDMRFTTPIRYPE	189
	Hs-FucA1	NWSPVSVWNWNSKDVG---PHRDLVGLGTALRKRNIRYGLHSL-LLEWPHFLYLDDKKN	194
		* * * * *	
GH29B	BinfantisFuc	-----GKA--YDDFYVGQLTELLTQYGPIFSVMLDGANGEGKNGKT--QYDWDYRYNV	194
	BtFuc1625	-----SPK--YNEEFIRQLTELLTNYGEVHEVWFDGANGEGPNGKK--QYDWDADFYKT	226
	BtFuc4136	-----SPR--YNKFFIRQLTELLTNYGEVHEVWFDGANGEGPNGKK--QYDWDTFVYET	226
	BtFuc2970	KEDSIAPSRF--LEETDNLKELATRMPTVKDFNFDGTWDASVKKNCWNTAHAEQMLKEL	253
GH29A	TmFuc	DLSYIRPNTYEMADYAYKQVMELVDLXL-PDVLKNDMGWPEKKGED-L--KILFAYYINK	245
	Hs-FucA1	G---FKTQHF-VSAKTMPELYDLVNSYK-PDLIWSGWECPDYWNSS--TNFLSWLYND	247
		* * * *	
GH29B	BinfantisFuc	IRSLQPDAVISVCGPDVVRWAGNEACHVRDNEWSVVPRLRSaelTMEKSQQEDDASFATT	254
	BtFuc1625	IQQLOPKAYMAIMGDDVVRWVGNEKGLGRETEWSATVLTTPG-----IYARSEENNKRLG	279
	BtFuc4136	IHRLOPKAYMAIMGDDIRWVGNESGLGRETEWSTTVLTPE-----IYARADKNNKRLG	279
	BtFuc2970	VF---GVAINSLRLADDKGRHFD-----SNGRLM--GD-----YESG	286
GH29A	TmFuc	HF---EGSVN-----DRWGVPHWD-----FKTAEY-----	267
	Hs-FucA1	SE-VKDEVVYN-----DRWGNCSG-----HHGGYY-----	272
		* * *	
GH29B	BinfantisFuc	VSSQDDDLGSRREAVAGYGDNVCVYPAEVDTSIR-FGWFIYHQSSE-DKVMASADQIFDLWLS	312
	BtFuc1625	VFSKAEDLGSRAMLE-KATELFWYFSEVDVSIR-FGWFIYHAED-SKVSKSLKHLADIYFQ	336
	BtFuc4136	INGQSNLDGSRKMLE-KATELFWYFSEVDVSIR-FGWFIYHKEED-NKVSKSLKHLADIYFQ	336
	BtFuc2970	YERRLPDPV-----KDLKVTQWDWEACMTIPENCNGYHKDWSLSYVKTPIEVIDRIHV	339
GH29A	TmFuc	---H---VN-----YPGDLPGYKWEFTRGIG-LSGYNRNNGPEHMLSVEQLVYTLVD	313
	Hs-FucA1	---NCEDKF-----KPQSLPDHKWEMCTSLDKFSYKCYRRDMALSDVTEESEIISELVQ	322
		* * *	

Figure 2 | Multiple sequence alignment of GH29 α -L-fucosidase homologs in the two subfamilies. The amino acids that are conserved in both subfamilies are noted with an asterisk underneath *. The residues known to interact with the C-6 position of L-fucose in TmFuc are denoted as orange rectangles and for BtFuc2970 as blue rectangles.

Among these α -L-fucosidases, the structure of two GH29A members, TmFuc and BtFuc2970, has been studied in-depth. Interestingly, in the structure of TmFuc a hydrophobic concavity, formed by aromatic amino acids F32, Y171, W222, and F290, surrounds the C-6 methyl group of L-fucose (marked in orange Figure 2).^{14,15} See Figure 3 for a schematic overview of the interaction shown for TmFuc and L-fucose as reported by Sulzenbacher and co-workers.¹⁵ A similar hydrophobic concavity was also seen in BtFuc2970. Structural analysis of BtFuc2970 showed that an additional O-6 hydroxyl moiety did not make any direct contacts with the protein; instead it occupied a gap within the active site surrounded by residues F64, W227, C308, and W316 (Blue in Figure 2).¹⁶ The idea of a hydrophobic cavity in or around the active site was already suggested in 1998 by Asano *et al.*¹⁷ They opted that this cavity might determine how C-6 substituents determined inhibition efficacy. As can be seen in Figure 2, not all amino acids known to form these hydrophobic cavities are within the conserved regions of the subset families. Thereby showing both the challenge as the potential of these C-6 substituents for achieving specificity among GH29 fucosidases.

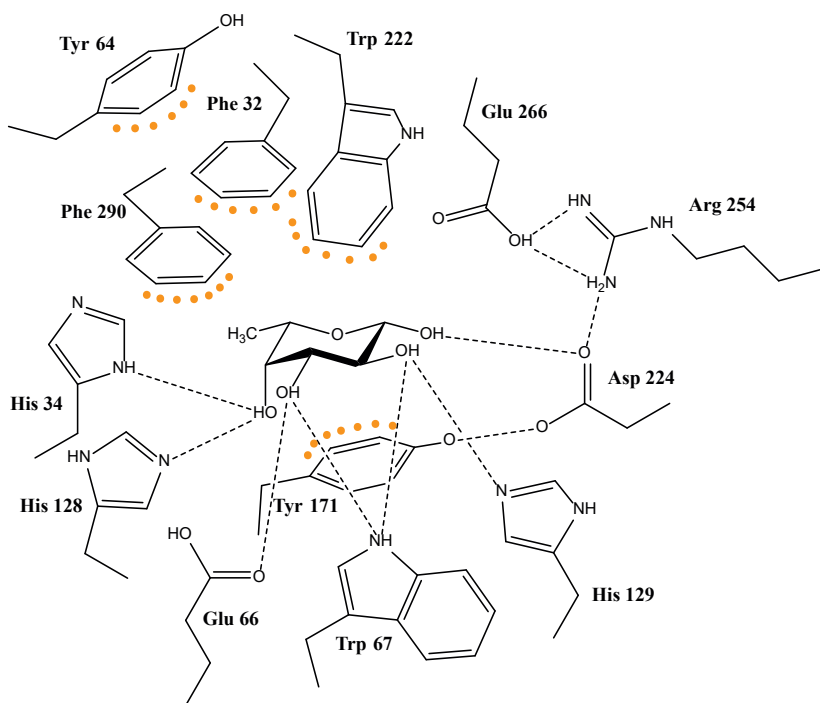
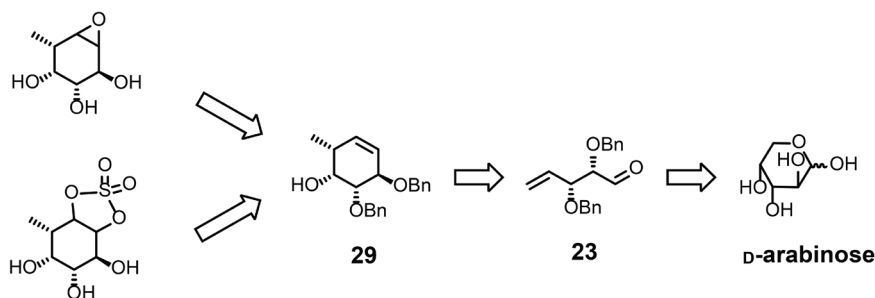


Figure 3 | Schematic overview of the interactions between TmFuc and L-fucose adapted from Sulzenbacher *et al.*¹⁵ The ligand is shown in boldface, hydrogen bonds are shown as dashed lines, and van der Waals contacts are shown as orange dotted lines.

At the onset of our studies, three distinct classes of covalent glycosidase inhibitors were known that could function as starting point in our endeavour towards higher specificity by C-6 modifications. These are the quinone methide glycosides, the 2-deoxy-2-fluoroglycosides, and the cyclophellitol-based epoxides. However, the quinone methide glycoside – although able to label recombinant purified enzymes and visualize fucosidase positive bacteria in co-cultures – seems to be an unsuitable lead as specific fucosidase inhibitor due to its broad and high reactivity (as described in **Chapter 4**). Previously reported studies on the other two classes of inhibitors seems more promising. 2-Deoxy-2-fluoroglycosides were first reported in 1987 by Withers and co-workers.¹⁸ The fluorine at the C-2 position destabilizes the oxacarbenium-like transition state, and thereby slows down the hydrolysis of the covalent glycosyl-enzyme intermediate. This leads to accumulation of the inhibitor-enzyme intermediate. However, a potential disadvantage of the 2-deoxy-2-fluoroglycosides is that the enzyme-inhibitor intermediate slowly hydrolyzes. The inhibition mechanism of the third class, cyclophellitol epoxides, comprises protonation of the epoxide in the active site, which is followed by nucleophilic attack of the carboxylate to give a covalent cyclophellitol-enzyme adduct that is not affected by hydrolysis.^{19,20} Recently, cyclophellitol cyclosulfates ABP analogues have also been reported by Artola *et al.* and showed promising inhibition of retaining glucosidases.²¹ In this section, we describe our efforts and future perspectives on the development of cyclophellitol epoxides and cyclophellitol cyclosulfate inhibitors and ABPs, that allow for C-6 modifications, as new potent research tools to study specific GH29 α -L-fucosidases.

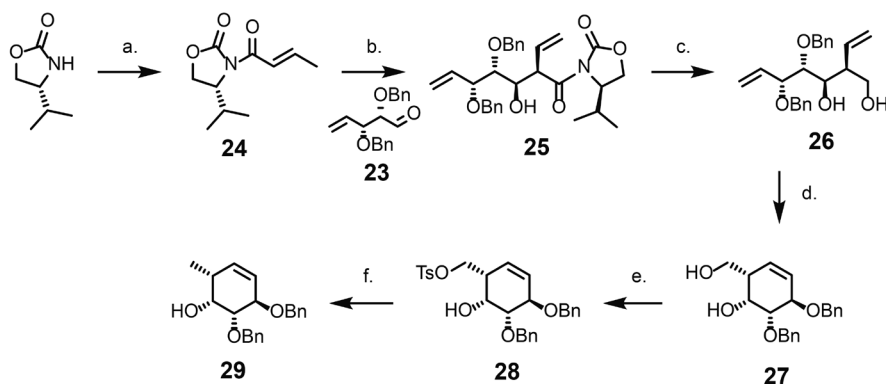
The inhibitors will be synthesized based on the same synthesis route as the L-fucopyranose-configured cyclophellitol aziridine, reported by Jiang *et al.*²² Scheme 4 depicts the retrosynthetic pathway for these targets, both can be synthesized from known intermediate **29**, which is derived from the building block **23**. Building block **23** can be synthesized from D-arabinose according to a protocol described by Hansen *et al.*²³



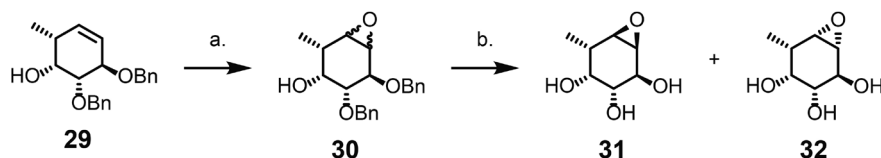
Scheme 4 | Retrosynthetic pathway of fucopyranose-configured cyclophellitol epoxide and cyclophellitol cyclosulfate inhibitors.

Cyclohexene derivative **29** was synthesized in five steps from building blocks **23** and **24** according to the route previously reported by Jiang *et al* (Scheme 5).²² The crucial step of this route is the aldol condensation of aldehyde **23** and the oxazolidinone **24**, catalyzed by dibutylboron triflate under strictly anhydrous conditions. Complete consumption of the activated oxazolidinone was observed on TLC when the reaction was left stirring overnight at -20°C, yielding compound **25** in 68% yield with the desired stereochemistry.

The cyclophellitol epoxide was successfully synthesized from compound **29** in two steps, as it is represented in scheme 6. Cyclohexene intermediate **29** was treated with *in situ* generated methyl(trifluoromethyl)dioxirane (TFD), resulting in a mixture of diastereomeric epoxides (**30**) in excellent yield. The benzyl protecting groups were removed by hydrogenation with palladium on carbon, affording a mixture of the α - and β -epoxides in a 6:1 ratio and a combined yield of 87%. The conformation of the α -L-fucopyranose-configured cyclophellitol epoxide **31** as the major isomer in this mixture was confirmed by NMR. A final separation to isolate **31** from the mixture is still required before it can be evaluated as a potential fucosidase inactivator.

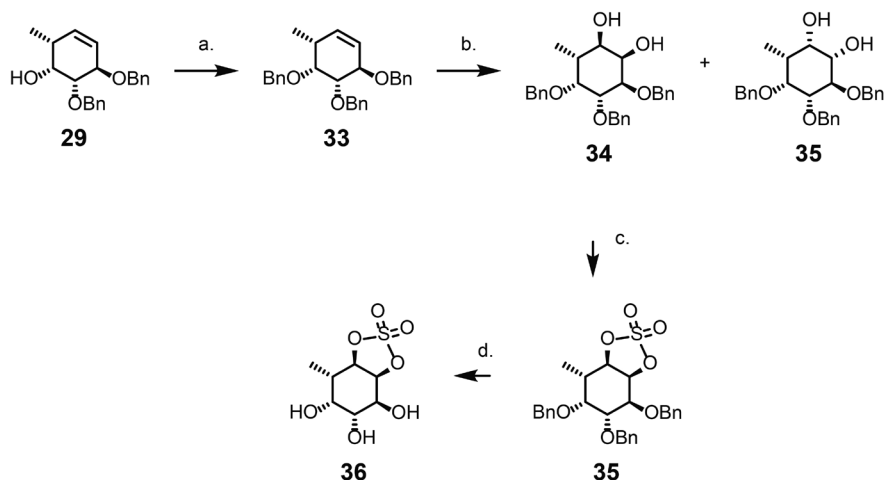


Scheme 5 | Synthesis of (1*R*,2*R*,5*R*,6*R*)-5,6-bis(benzyloxy)-2-methylcyclohex-3-en-1-ol (**29**). *Reagents and conditions*: (a) Crotonyl chloride, *n*-BuLi, THF, 30 min, -78°C to rt, 99%; (b) DBBT, Et₃N, DCM, 24h, -78°C to 0°C, 68%; (c) LiBH₄, THF/H₂O, 2h, rt, 99%; (d) Grubbs 2nd generation, DCM, 24h, 40°C, 62%; (e) TsCl, pyridine, DCM, 24h, rt, 76%; (f) LiAlH₄, THF, 3h, rt, 93%.



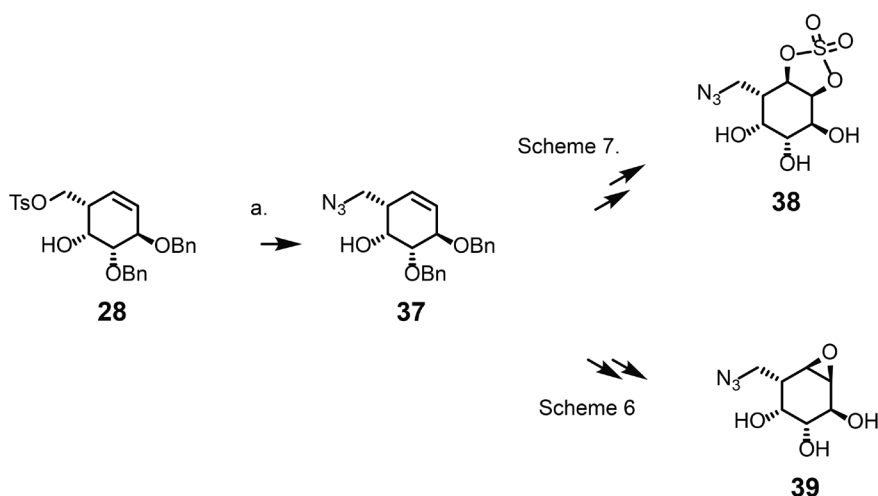
Scheme 6 | Synthesis of fucopyranose-configured cyclophellitol epoxides (**31**) and (**32**). *Reagents and conditions*: (a) Oxone, NaHCO₃, CH₃COCF₃, MeCN/EDTA, 1½ h, 0°C, 99%; (b) H₂, Pd/C, MeOH/EtOAc, 4h, rt, 87%.

The synthesis of the L-fucopyranose-configured cyclophellitol cyclosulfate started with the benzylation of the free hydroxyl group of intermediate **29** with benzyl bromide, sodium hydride and a catalytic amount of tetrabutylammonium iodide (TBAI), followed by oxidation with RuO_4 , *in situ* generated by $\text{RuCl}_3 \cdot x\text{H}_2\text{O}$ and NaIO_4 , which dihydroxylates the double bond of **33** (Scheme 7). The major product was the β -L-fucopyranose-configured diol **35** with 16% yield. The conformation was confirmed with NMR. The minor spots observed on TLC were analyzed but the α -conformation of the diol **34** could not be confirmed by either NMR or TLC-MS. To potentially afford the desired α -conformation and increase the yield of the oxidation of alkene **33**, a catalytic amount of OsO_4 , which is reported as a stronger oxidant for syn-diols, could be used. To continue this route, the diol **34** could be treated with thionyl chloride to form the cyclic sulfite and subsequently treated with *in situ* synthesized RuO_4 to oxidize the sulfite to sulfate, yielding the protected α -fucopyranose-configured cyclosulfate **35**. To achieve deprotection, compound **35** will be hydrogenated with palladium on carbon.



Scheme 7 | Synthesis of L-fucopyranose-configured cyclophellitol cyclosulfate (**36**). *Reagents and conditions:* (a) TBAI, BnBr, NaH, DMF, 24h, rt, 68%; (b) $\text{RuCl}_3 \cdot x\text{H}_2\text{O}$, NaIO_4 , EtOAc/ACN, H_2O , 2½ h, 0°C, 16%; (c) (i) SOCl_2 , Et_3N , DCM, 32h, 0°C-rt; (ii) $\text{RuCl}_3 \cdot x\text{H}_2\text{O}$, NaIO_4 , EtOAc/ACN, H_2O , 24h, 0°C, 16%; (d) H_2 , Pd/C, MeOH/EtOAc, 4h, rt.

After establishing the inhibitory efficacy of the newly synthesized inhibitors, the next step would be to transform these inhibitors into ABPs. A similar approach to the described synthesis of **Chapter 3** will be used, in which the C-6 position will be used for the introduction of an azide moiety allowing facile attachment of other substituents. Intermediate tosylated compound **28** will first be treated with sodium azide and DMF to install the azide on the primary position, as it is depicted in scheme 8. Finally, the synthesis should be continued as previously described in scheme 6 and 7 to afford the cyclophellitol epoxide and the cyclophellitol cyclosulfate ABPs. The combination of first principles molecular dynamics and structural approaches allows exploitation to design and evaluate specific C-6 modified ABPs and holds potential for targeting specific GH29 α -L-fucosidases.



Scheme 8 | Proposed synthesis for ABPs of α -L-fucopyranose-configured cyclophellitol epoxide and cyclophellitol cyclosulfate. *Reagents and conditions:* (a) NaN_3 , DMF, 3h, 90°C .

6.3 Concluding Remarks

In **chapter 2** and **3** of this thesis it was highlighted the complexity of the human-gut microbiota interface at the level of its glycans. It could be concluded that membrane bound *O*-glycosylation of the epithelial cells is a dynamic process dependent on several factors, such as the differentiation state of the cells as the presence of bacterial metabolites. In parallel to this it was also seen that fucosidases secreted by residing commensal bacteria can be abused by pathogens for cross feeding mechanisms. These studies underline the importance and strength of designing high-throughput glycomics methods and new molecular tools such as ABPs that allow for future studies towards understanding what is going on in complex gut microbiota samples at the molecular level when focussing on fucosylation and fucosidases. With respect to what is reported here, the next step with respect to *O*-glycomics would be to gain insight into the dynamics during differentiation of secreted (MUC2) mucin *O*-glycosylation.

What to do next with the in this thesis reported novel fucosidase probes? The fucosidase-based ABP presented in **chapter 4** might encounter some hurdles if it were to be applied *in vivo* due to the probable increased hydrolysis of the covalent intermediate it forms under the basic conditions of the gut. However, an application of this probe as a diagnostic tool on pre-treated microbiota samples, where there is more control over certain parameters, is certainly viable. The probe presented in **chapter 5** can be further developed for this same application and is unique in its ability to detect and label both retaining and inverting fucosidases. However, more complex microbiota samples might be challenging due to non-specific labeling that this class of probe is prone to. In a pilot experiment where a bacterium not capable of expressing fucosidases was incubated in the presence of the probe with and without additionally added recombinant fucosidases, cross-labelling could be observed. It would be interesting to investigate if the addition of additives, such as Tris buffers that have been reported to efficiently capture quinone methides, can result in less non-specific binding.²⁴

Additionally, there is ample room to explore modifications of the ABP design to improve and tune their activity and selectivity. Over the last years multidisciplinary approaches have been reported that combine molecular dynamics and structural approaches into the design towards very specific ABPs.²⁵ Future studies based on the here described ABP designs combined with the insights from molecular dynamics and structural approaches is therefore a promising new direction to develop new highly specific probes to expand the molecular toolbox to study fucosidases. In the future this toolbox might provide excellent opportunities to study the direct effect of specific fucosidases in complex gut microbiota samples.

Experimental

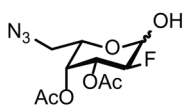
General methods en materials

All chemicals were obtained from commercial suppliers and used as received unless stated otherwise. Organic solvents for reactions were dried for at least 2 days over molecular sieves (3 or 4 Å). ^1H and ^{13}C NMR spectra were recorded on either an Agilent 400 instrument (400 and 101 MHz) or a Bruker Avance Neo 600 spectrometer (600 and 125 MHz). Chemical shifts are reported in δ values relative to TMS and J coupling constants are reported in Hertz (Hz). Silica column chromatography was performed using silica gel SiliaFlash P60 (SiliCycle, Canada, 40–63 μm , 239–400 mesh). TLC analysis was conducted on SiliaPlate TLC Aluminium Backed TLC F254 (SiliCycle) with examination under UV light (254 nm) where applicable, and with 5% sulfuric acid in ethanol or with ceric ammonium molybdate, followed by heating.

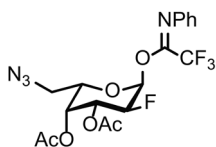
General method A – fluorination of glucals: To a solution of the glucal (1 eq.) in DMF was added H_2O and F-TEDA-BF₄ (Selectfluor) (1.5 eq.). The reaction mixture was stirred at room temperature for 32 h, diluted with EtOAc, and washed with water. The organic layer was dried over MgSO_4 and concentrated. The residue was purified by silica gel column chromatography to provide the 2-deoxy-2-fluoro- glycoside products.

General method B – Deacetylation under Zemplén conditions: To a solution of acetylated compound in MeOH was added NaOMe and stirred for 2h. The reaction was quenched with Amberlite-H⁺ IR-120 till pH \leq 7, filtered and concentrated *in vacuo* yielding deacetylated compounds.

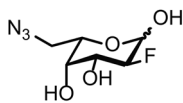
General method C – Installation of the anomeric N-phenyl trifluoro imidate: To a 0 °C cooled solution of 2-deoxy-2-fluoro glycosides (1 eq.) in acetone was added *N*-phenyl trifluoroacetimidoyl chloride (2 eq.) and K_2CO_3 (2 eq.). The reaction was gradually warmed to rt and stirred for 2h. The solids were filtered, and the filtrate was concentrated *in vacuo*. The residue was purified by silica gel column chromatography to provide the 2-deoxy-2-fluoro- glycosyl imidate products.

3,4-di-O-acetyl-6-azido-2-fluoro- α/β -L-galactoside 6

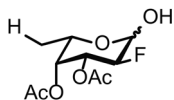
Compound **5** was subjected to general procedure A to produce **6** (31 mg, 0.11 mmol) in 73 % yield after silica gel column purification (9:1 PE/EtOAc). ^1H NMR (400 MHz, CDCl_3) α/β mixture δ 5.53 (d, $J = 3.7$ Hz, 3H), 5.49 (d, $J = 3.1$ Hz, 1H), 5.43 (d, $J = 1.1$ Hz, 3H), 5.38 – 5.20 (m, 3H), 5.10 (ddd, $J = 13.1, 9.8, 3.6$ Hz, 1H), 4.75 (ddd, $J = 50.2, 9.6, 3.7$ Hz, 3H), 4.54 (dd, $J = 9.9, 7.6$ Hz, 0.5H), 4.44 – 4.35 (m, 3.5H), 4.33 – 4.27 (m, 1H), 3.48 (dd, $J = 12.9, 7.9$ Hz, 1H), 3.39 (dd, $J = 12.8, 8.1$ Hz, 3H), 3.20 (ddd, $J = 12.9, 9.6, 4.6$ Hz, 4H), 2.14 (s, 3H), 2.13 (s, 9H), 2.04 (s, 3H), 2.02 (s, 9H). ^{13}C NMR (101 MHz, CDCl_3) α/β mixture δ 170.50, 170.15, 92.16, 90.53, 89.88, 88.02, 86.58, 84.70, 70.86, 69.48, 68.60, 68.04, 67.94, 50.55, 50.40, 20.82, 20.53.

3,4-di-O-acetyl-6-azido-2-fluoro- α -L-galactosyl imidate 7

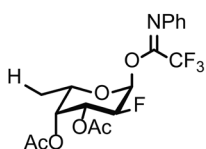
Compound **6** was subjected to general procedure C to produce **7** (32 mg, 0.069 mmol) in 63% yield after silica gel column purification (9:1 PE/EtOAc). ^1H NMR (400 MHz, CD_3CN) δ 7.36 (dd, $J = 8.4, 7.5$ Hz, 2H), 7.20 – 7.13 (m, 1H), 6.90 (d, $J = 7.2$ Hz, 2H), 6.60 (s, 1H), 5.47 (td, $J = 3.5, 1.2$ Hz, 1H), 5.41 (td, $J = 10.9, 3.5$ Hz, 1H), 5.05 (ddd, $J = 48.1, 10.0, 3.7$ Hz, 1H), 4.36 (dd, $J = 8.8, 4.3$ Hz, 1H), 3.56 – 3.18 (m, 2H), 2.13 (s, 3H), 2.02 (s, 3H). ^{13}C NMR (101 MHz, CD_3CN) δ 170.11, 169.77, 128.95, 124.53, 119.06, 85.46, 83.57, 70.46, 68.40, 67.96, 49.83, 19.71. HR MS (m/z) calcd for $\text{C}_{18}\text{H}_{18}\text{F}_4\text{N}_4\text{O}_6$ [$\text{M} + \text{Na}$] $^+$, 485,3492; found, 485,1056.

6-azido-2,6-dideoxy-2-fluoro- α -L-galactosyl 8

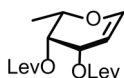
Compound **6** was subjected to Zemplén saponification according to general protocol B to produce **8** (31 mg, 0.15 mmol) in quantitative yield. ^1H NMR (400 MHz, CD_3OD) δ 5.47 (d, $J = 4.0$ Hz, 1H), 4.82 (dd, $J = 7.2, 2.9$ Hz, 1H), 4.18 – 3.95 (m, 1H), 3.81 (d, $J = 4.4$ Hz, 1H), 3.75 (d, $J = 3.2$ Hz, 1H), 3.48 (d, $J = 7.3$ Hz, 1H), 3.43 (d, $J = 5.9$ Hz, 1H). ^{13}C NMR (101 MHz, CD_3OD) δ 99.87, 96.87, 95.13, 94.69, 71.22, 70.78, 51.50.

3,4-di-O-acetyl-2-fluoro- α/β -L-fucoside 11

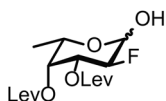
3,4-di-O-acetyl-L-fucal was subjected to general procedure A to produce **11** as an α/β mixture (165 mg, 0.66 mmol) in 53 % yield after silica gel column purification (8:2 PE/EtOAc). Spectroscopic data in accordance with known literature data.²⁶

3,4-di-O-acetyl-2-fluoro- α/β -L-fucosyl imidate **12**

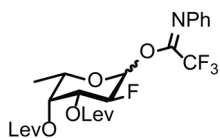
Compound **11** was subjected to general procedure C to produce **12** (169 mg, 0.40 mmol) in 87% yield after silica gel column purification (9:1 PE/EtOAc). ^1H NMR (400 MHz, CDCl_3) α/β mixture δ 7.30 (d, J = 3.9 Hz, 6H), 7.15 – 7.08 (m, 3H), 6.83 (d, J = 6.1 Hz, 6H), 6.58 (s, 1H), 5.47 (t, J = 4.1 Hz, 1H), 5.40 (d, J = 3.6 Hz, 2H), 5.32 (d, J = 3.7 Hz, 1H), 5.26 (s, 2H), 5.10 (ddd, J = 13.2, 9.8, 3.5 Hz, 2H), 5.01 – 4.82 (m, 1H), 4.82 – 4.62 (m, 2H), 4.46 – 4.37 (m, 1H), 4.29 (d, J = 7.0 Hz, 2H), 1.27 – 1.01 (m, 9H). ^{13}C NMR (101 MHz, CDCl_3) α/β mixture δ 129.04, 124.93, 119.43, 90.75, 87.22, 85.77, 83.74, 71.87, 71.52, 71.00, 70.42, 68.74, 67.47, 64.86, 20.72, 15.49. HR MS (m/z) calcd for $\text{C}_{18}\text{H}_{19}\text{F}_4\text{NO}_6$ [$\text{M} + \text{Na}$] $^+$, 444,3292; found, 444,1033.

3,4-di-O-levulinoyl-fucal **15**

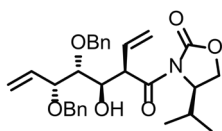
Di-acetylated L-fucal **10** was deprotected using Zemplén conditions according to general procedure B to yield L-fucal **14** (200 mg, 1.54 mmol) in quantitative yields. The obtained crude residue was dissolved in dry DCM (10 mL) and levulinic acid (0.38 mL, 3.6 mmol, 2.4 eq.), EDC (0.88 mg, 4.5 mmol, 3 eq.) and DMAP (catalytic amount) were added. The resulting mixture was stirred under argon for 32h. After which the reaction mixture was diluted with DCM and washed with H_2O and sat. aq. NaHCO_3 . The organic layer was separated, dried over MgSO_4 , and concentrated *in vacuo*. The residue was purified by silica gel column purification (7:3 EtOAc/PE) to afford **15** (430 mg, 1.3 mmol) as a yellow syrup in 84% yield. ^1H NMR (400 MHz, CDCl_3) δ 6.44 (dd, J = 6.3, 1.8 Hz, 1H), 5.53 (dq, J = 3.4, 1.1 Hz, 1H), 5.27 – 5.25 (m, 1H), 4.24 – 4.16 (m, 1H), 4.15 – 4.04 (m, 1H), 2.82 – 2.72 (m, 4H), 2.71 – 2.56 (m, 4H), 1.25 (td, J = 7.1, 1.1 Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 206.56, 206.22, 172.27, 171.11, 146.05, 98.21, 66.35, 65.30, 60.36, 29.84, 27.89, 21.02.

3,4-di-O-levulinoyl-2-fluoro- α/β -L-fucoside **16**

Compound **15** was subjected to general procedure A to produce **16** as an α/β mixture (235 mg, 0.65 mmol) in 49% yield after silica gel column purification (7:3 PE/EtOAc). ^1H NMR (400 MHz, CDCl_3) α/β mixture δ 5.46 (s, 1H), 5.46 – 5.37 (m, 1H), 5.34 – 5.29 (m, 1H), 5.27 – 5.23 (m, 1H), 5.07 (dddd, J = 13.3, 9.9, 3.7, 1.4 Hz, 1H), 4.87 – 4.79 (m, 2H), 4.69 (ddd, J = 10.2, 3.8, 1.3 Hz, 1H), 4.55 – 4.34 (m, 1H), 3.91 – 3.81 (m, 1H), 2.81 – 2.63 (m, 8H), 1.20 (d, J = 6.5 Hz, 3H), 1.13 (d, J = 6.6 Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) α/β mixture δ 206.61, 206.31, 172.31, 172.06, 94.59, 90.58, 90.23, 88.37, 86.85, 84.97, 71.88, 71.76, 70.96, 69.45, 68.52, 64.64, 29.79, 27.64, 15.94, 15.72.

3,4-di-O-levulinoyl-2-fluoro- α/β -L-fucosyl imidate 17

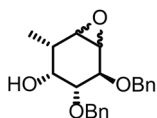
Compound **16** was subjected to general procedure C to produce **17** (180 mg, 0.34 mmol) in 68% yield after silica gel column purification (7:3 PE/EtOAc). ^1H NMR (400 MHz, CDCl_3) α/β mixture δ 7.34 – 7.27 (m, 2H), 7.16 – 7.08 (m, 1H), 6.90 – 6.79 (m, 2H), 6.59 (s, 0.5H), 5.84 (s, 0.5H), 5.49 – 5.35 (m, 0.5H), 5.28 (s, 0.5H), 5.10 (s, 0.5H), 4.93 (dd, J = 49.0, 9.7 Hz, 0.5H), 4.73 (d, J = 51.8 Hz, 0.5H), 4.29 (d, J = 6.7 Hz, 0.5H), 4.24 (dd, J = 5.8, 4.3 Hz, 0.5H), 2.83 – 2.49 (m, 8H), 1.24 – 1.15 (m, 3H). ^{13}C NMR (101 MHz, CDCl_3) α/β mixture δ 206.58, 206.26, 172.35, 171.97, 128.91, 124.59, 119.64, 87.83, 86.39, 86.03, 83.60, 72.16, 71.08, 70.72, 68.91, 67.83, 27.92, 16.12.

(R)-3-((2R,3R,4R,5R)-4,5-bis(benzyloxy)-3-hydroxy-2-vinylhept-6-enoyl)-4-isopropylloxazolidin-2-one 25

Et_3N was freshly distilled (under N_2 atmosphere and KOH) and dried over KOH. Oxazolidinone and aldehyde were dissolved in anhydrous DCM and dried over freshly activated molecular sieves (4\AA) and argon atmosphere for at least 1 hour before starting the reaction. Oxazolidinone (**24**) (30 mg, 0.15 mmol, 1.0 eq.), dissolved in anhydrous DCM (1 ml), was cooled to -78°C and DBBT (0.18 ml, 0.18 mmol, 1.2 eq.) was added with a syringe under argon atmosphere (color turned to orange). After stirring for 15 minutes, Et_3N (0.03 ml, 0.21 mmol, 1.4 eq.) was added dropwise, causing the dark color to fade. The mixture was stirred for 1 hour at -78°C and for 15 minutes at 0°C , resulting in a bright yellow color. The reaction was cooled back to -78°C and a solution of the aldehyde (**23**) (43 mg, 0.2 mmol, 1.0 eq.) in anhydrous DCM was added dropwise and the temperature was slowly raised to -20°C for 1 hour and the reaction was stirred at this temperature overnight. After completion, the orange solution was quenched with PBS (1 ml) and 30% H_2O_2 was added until the internal temperature remained unchanged and below 5°C . The mixture was stirred for 45 minutes at rt and it was poured over saturated aqueous NaHCO_3 , extracted with DCM, dried over Na_2SO_4 , filtered, and concentrated under reduced pressure. The residue was purified with silica gel column chromatography (9:1 > 8:2, PE/EtOAc) giving compound (**25**) as a colorless oil (50 mg, 0.1 mmol, 68%). ^1H NMR (600 MHz, CDCl_3) δ 7.26 (td, J = 8.6, 2.1 Hz, 7H), 7.24 – 7.17 (m, 3H), 6.04 – 5.95 (m, 1H), 5.92 – 5.83 (m, 1H), 5.38 – 5.31 (m, 3H), 5.22 (dd, J = 10.2, 1.5 Hz, 1H), 4.97 (dd, J = 8.9, 7.1 Hz, 1H), 4.66 – 4.56 (m, 2H), 4.47 (d, J = 11.2 Hz, 1H), 4.44 – 4.35 (m, 2H), 4.23 (dd, J = 7.3, 3.9 Hz, 1H), 3.97 (dt, J = 8.7, 3.4 Hz, 1H), 3.77 (dd, J = 9.0, 3.1 Hz, 1H), 3.52 (dd, J = 8.3, 3.9 Hz, 1H), 3.36 (d, J = 2.7 Hz, 1H), 3.22 (t, J = 8.8 Hz, 1H), 2.15 (pd, J = 7.0, 3.8 Hz, 1H), 0.69 (dd, J = 25.6, 7.0 Hz, 6H). ^{13}C NMR (101 MHz, CDCl_3) δ 172.50, 153.49, 137.99, 137.70, 134.45, 133.48, 120.43, 119.15, 81.71, 79.86, 72.96, 71.13, 70.65, 62.37, 57.95, 50.09, 27.96, 17.75, 14.40.

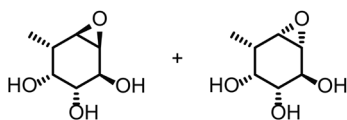
Data in accordance with previous literature.²²

(1R,2S,3R,4R,5S,6R)-4,5-bis(benzyloxy)-2-methyl-7-oxabicyclo[4.1.0]heptan-3-ol **30**



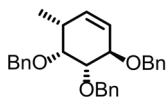
Alkene **29** (100 mg, 0.3 mmol, 1.0 eq.) was dissolved in a 2:1 mixture of MeCN/EDTA 0.1M (3 ml) and 1,1,1-Trifluoroacetone (0.4 ml, 4.65 mmol, 15.0 eq.) was added at 0°C. The mixture of solids NaHCO₃ (182 mg, 2.17 mmol, 7.0 eq.) and Oxone (236 mg, 1.55 mmol, 5.0 eq.) was added over 5 minutes, maintaining a stable pH 7. The reaction mixture was stirred at 0°C for 1½ hours. After completion, the reaction was quenched with H₂O, extracted with EtOAc, washed with brine, dried over Na₂SO₄, filtered and concentrated under reduced pressure. The crude product was purified with silica gel column chromatography (9:1 > 7:3 PE/EtOAc) to afford a mixture of isomers of epoxide (**30**) as a colorless oil (105 mg, 0.3 mmol, 100%).

(1S,2S,3R,4R,5S,6R)-5-methyl-7-oxabicyclo[4.1.0]heptane-2,3,4-triol **31** and **32**



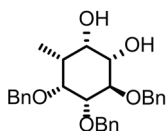
Compound **30** (105 mg, 0.3 mmol, 1.0 eq.) was dissolved in a 2:1 mixture of MeOH/EtOAc (25 ml) and the flask was flushed with N₂ for 5 minutes. Palladium on carbon (6.4 mg, 0.06 mmol, 0.2 eq.) was added and a balloon of H₂ was placed and bubbled the mixture. After 5 minutes, the stop was placed on the outlet and the needle was lifted slightly out of the solution. The reaction mixture was stirred at rt for 4 hours. After completion, the mixture was filtered through a Celite plug and concentrated under reduced pressure. The crude product was purified with silica gel column chromatography (9:1 > 8:2 DCM/MeOH) to afford the α (**31**) and the β diastereomers (**32**) in a 6:1 ratio as a white solid (41.7 mg, 0.26 mmol, 87%). ¹H NMR (600 MHz, CD₃OD) δ 4.03 (dd, *J* = 8.7, 2.4 Hz, 1H), 3.58 (dt, *J* = 3.7, 1.9 Hz, 1H), 3.38 (dd, *J* = 8.7, 1.9 Hz, 1H), 3.27 (dd, *J* = 4.0, 2.4 Hz, 1H), 2.83 (dd, *J* = 4.0, 1.8 Hz, 1H), 2.00 (dd, *J* = 7.7, 3.5 Hz, 1H), 1.12 (d, *J* = 7.8 Hz, 3H). ¹³C NMR (151 MHz, CD₃OD) δ 74.95, 73.84, 70.46, 58.55, 58.25, 36.29, 15.13.

(((1R,2S,3R,6R)-6-methylcyclohex-4-ene-1,2,3-triyl)tris(oxy)) tris(methylene) tribenzene
33



Alkene **29** (50 mg, 0.15 mmol, 1.0 eq.) was dissolved in DMF (10 ml) at 0°C and TBAI (5.5 mg, 0.015 mmol, 0.1 eq.), BnBr (0.09 ml, 0.75 mmol, 5.0 eq.) and NaH (60% dispersion in mineral oil, 30 mg, 0.75 mmol, 5.0 eq.) were added. The reaction mixture was stirred at rt overnight. After completion, the reaction was quenched with MeOH at 0°C, diluted with water and subsequently extracted with EtOAc. The organic layer was washed with brine, dried over Na₂SO₄, filtered and concentrated under reduced pressure. The crude product was purified with silica gel column chromatography (9:1 PE/EtOAc) to afford compound (**33**) as a colorless oil (42 mg, 0.1 mmol, 68%). ¹H NMR (600 MHz, CDCl₃) δ 7.37 – 7.23 (m, 15H), 5.70 (dt, *J* = 10.0, 2.6 Hz, 1H), 5.45 (d, *J* = 10.2 Hz, 1H), 4.93 – 4.66 (m, 6H), 4.43 (dt, *J* = 7.7, 2.3 Hz, 1H), 3.89 (dd, *J* = 3.6, 1.7 Hz, 1H), 3.76 (dd, *J* = 7.6, 1.7 Hz, 1H), 1.07 (d, *J* = 7.4 Hz, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 139.03, 138.87, 135.21, 128.60, 128.45, 128.32, 128.30, 128.26, 128.13, 127.99, 127.75, 127.49, 127.41, 127.36, 83.21, 77.90, 77.48, 72.31, 72.16, 35.76, 16.54.

(1S,2S,3R,4R,5R,6R)-3,4,5-tris(benzyloxy)-6-methylcyclohexane-1,2-diol **35**



Alkene **33** (42 mg, 0.1 mmol, 1.0 eq.) was dissolved in a 1:1 mixture of EtOAc/ACN (5 ml) and it was cooled to 0°C. Catalytic amount of RuCl₃·xH₂O (1 mg, 0.005 mmol, 0.05 eq.) and NaIO₄ (32 mg, 0.15 mmol, 1.5 eq.) were dissolved in H₂O (1.1 ml) and were added to the reaction mixture, which was subsequently stirred for 2½ h at 0°C. After completion, the reaction was quenched with 10% aqueous Na₂S₂O₃ and the aqueous layer was separated and extracted three times with EtOAc. The combined organic layers were washed with brine, dried over Na₂SO₄, filtered and concentrated under reduced pressure. The crude product was purified with silica gel column chromatography (9:1 > 4:6 PE/EtOAc) to afford the β enantiomer (**35**) as a colorless oil (7.2 mg, 0.016 mmol, 16%). ¹H NMR (600 MHz, CDCl₃) δ 7.35 (m, *J* = 31.0, 14.6, 7.6 Hz, 15H), 5.03 (d, *J* = 11.0 Hz, 1H), 4.94 (s, 2H), 4.81 (d, *J* = 11.7 Hz, 1H), 4.74 (d, *J* = 11.8 Hz, 1H), 4.58 (d, *J* = 11.1 Hz, 1H), 3.94 (t, *J* = 9.5 Hz, 1H), 3.87 (s, 1H), 3.72 (d, *J* = 10.2 Hz, 1H), 3.51 (d, *J* = 10.0 Hz, 1H), 3.45 (dd, *J* = 9.7, 2.5 Hz, 1H), 1.22 (d, *J* = 7.1 Hz, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 128.48, 128.37, 128.16, 128.12, 128.04, 127.61, 127.56, 83.87, 81.79, 80.59, 75.93, 75.86, 75.82, 74.90, 73.27, 34.71, 14.23.

Acknowledgements

Stefania Douka is thanked for the synthesis of the cyclophellitol based inhibitors and the fruitful discussions on the problems encountered in our attempts to obtain the α -configured fucopyranoside epoxides and cyclosulfates. Enrico Verpalen and Arwin Brouwer are kindly acknowledged for their fruitful discussions on the synthesis of the 2-deoxy-2-fluoro galactopyranosyl imidate.

References

1. Hsu YL, Nandakumar M, Lai HY, Chou TC, Chu CY, Lin CH, Lo LC. *J Org Chem*. 2015; 80: 8458–8463.
2. Rempel BP, Withers SG. *Glycobiology*. 2008; 18, 570–586.
3. Jeroen C Codée CD, Overkleeft HS, C Walvoort MT, Kallemeijn WW, Willems LI, Witte MD, F G Aerts JM, van der Marel GA, C Codeé JD. *Chem Commun Chem Commun*. 2012; 4848841: 10357–10456.
4. Jensen HH, Bols M. *Org Lett*. 2003; 5: 3419–3421.
5. Lucas-Lopez C, Murphy N, Zhu X. *European J Org Chem*. 2008; 2008: 4401–4404.
6. Demchenko A V., Pornsuriyasak P, De Meo C, Malysheva NN. *Angew Chemie - Int Ed*. 2004; 43: 3069–3072.
7. Bernardes GJL, Gamblin DP, Davis BG. *Angew Chemie - Int Ed*. 2006; 45: 4007–4011.
8. Zhu X. *Tetrahedron Lett*. 2006; 47: 7935–7938.
9. Panmontha W, Amarinthnukrowh P, Damrongphol P, Desudchit T, Suphapeetiporn K, Shotelersuk V. *Genet Mol Res*. 2016; 15.
10. Stütz AE, Wrodnigg TM. In *Advances in Carbohydrate Chemistry and Biochemistry* Academic Press Inc., 2016; Vol. 73, pp 225–302.
11. Wu HJ, Ho CW, Ko TP, Popat SD, Lin CH, Wang AH -J. *Angew Chemie - Int Ed*. 2010; 49: 337–340.
12. Lammerts Van Bueren A, Popat SD, Lin CH, Davies GJ. *ChemBioChem*. 2010; 11: 1971–1974.
13. Sakurama H, Tsutsumi E, Ashida H, Katayama T, Yamamoto K, Kumagai H. *Biosci Biotechnol Biochem*. 2012; 76: 1022–1024.
14. Bishnoi R, Mahajan S, Ramya TNC. *Glycobiology*. 2018; 28: 860–875.
15. Sulzenbacher G, Bignon C, Nishimura T, Tarling CA, Withers SG, Henrissat B, Bourne Y. *J Biol Chem*. 2004; 279: 13119–13128.
16. Van Bueren AL, Ardèvol A, Fayers-Kerr J, Luo B, Zhang Y, Sollogoub M, Blériot Y, Rovira C, Davies GJ. *J Am Chem Soc*. 2010; 132: 1804–1806.
17. Asano N, Nishida M, Kato A, Kizu H, Matsui K, Shimada Y, Itoh T, Baba M, Watson AA, Nash RJ, De Lilley PMQ, Watkin DJ, Fleet GWJ. *J Med Chem*. 1998; 41: 2565–2571.
18. Withers SG, Street IP, Bird P, Dolphin DH. *J Am Chem Soc*. 1987; 109: 7530–7531.
19. Withers SG, Umezawa K. *Biochem Biophys Res Commun*. 1991; 177: 532–537.
20. Atsumi S, Iinuma H, Naganawa H, Nakamura H, Takeuchi T, Umezawa K, Iitaka Y. *J Antibiot (Tokyo)*. 1990; 43: 49–53.
21. Artola M, Wu L, Ferraz MJ, Kuo CL, Raich L, Breen IZ, Offen WA, Codée JDC, Van Der Marel GA, Rovira C, Aerts JMFG, Davies GJ, Overkleeft HS. *ACS Cent Sci*. 2017; 3: 784–793.
22. Jiang J, Kallemeijn WW, Wright DW, Van Den Nieuwendijk AMCHCH, Rohde VC, Folch EC, Van Den Elst H, Florea BI, Scheij S, Donker-Koopman WE, Verhoek M, Li N, Schürmann M, Mink D, Boot RG, Codée JDCC, Van Der Marel GA, Davies GJ, Aerts JMFGFG, Overkleeft HS, Coco Rohde V, Colomina Folch E, Van Den Elst H, Florea BI, Scheij S, Donker-Koopman WE, Verhoek M, Li N, Schürmann M, Mink D, Boot RG, Codée JDCC, van den Marel GA, Davies GJ, Aerts JMFGFG, Overkleeft HS. *Chem Sci*. 2015; 6: 2782–2789.
23. Hansen FG, Bundgaard E, Madsen R. *J Org Chem*. 2005; 70: 10139–10142.

24. Bruins JJ, Albada B, van Delft F. *Chem. - A Eur. J.*, 2018, 24, 4749–4756.
25. McGregor NGS, Artola M, Nin-Hill A, Linzel D, Haon M, Reijngoud J, Ram A, Rosso MN, Van Der Marel GA, Codeé JDC, Van Wezel GP, Berrin JG, Rovira C, Overkleeft HS, Davies GJ. *J Am Chem Soc.* 2020; 142: 4648–4662.
26. Burkart MD, Zhang Z, Hung SC, Wong CH. *J Am Chem Soc.* 1997; 119: 11743–11746.

Appendix

Nederlandse samenvatting

Curriculum Vitae

List of Publications

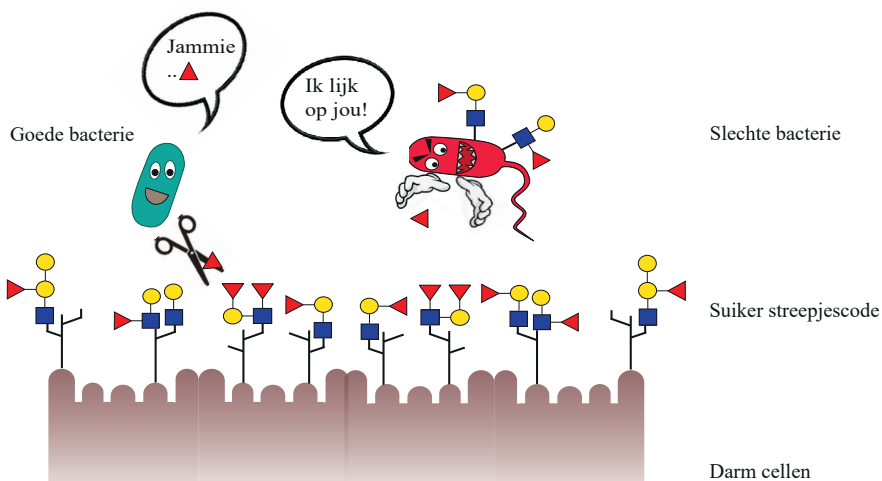
Acknowledgement

Nederlandse samenvatting

Ondanks dat bacteriën van vroeger uit een slechte naam hebben, zijn ze zeker niet altijd de slechteriken maar beschermen ons juist ook tegen mogelijke infecties. In onze darmen zijn bijvoorbeeld enorme hoeveelheden bacteriën aanwezig die een belangrijke rol spelen in het behoud van onze gezondheid.

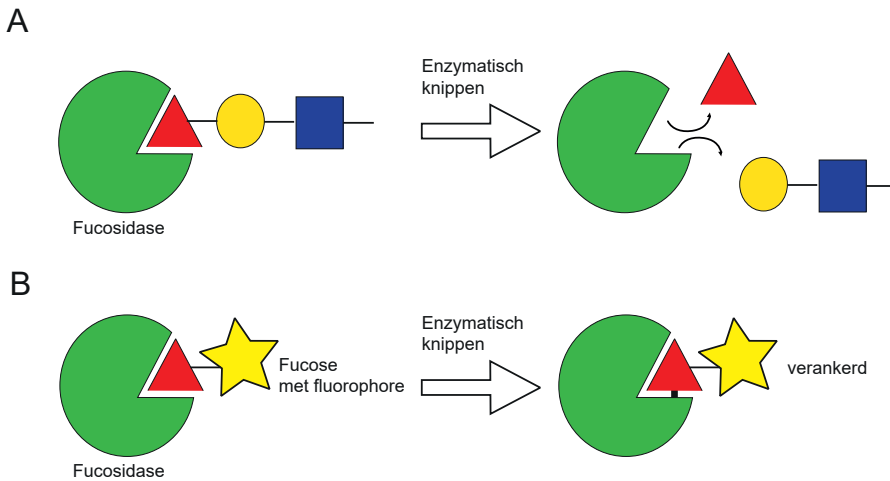
Onderzoekers proberen deze darmbewoners in kaart te brengen om zo meer te weten te komen over hun effect op het menselijk lichaam. Naast het achterhalen van welke bacteriën er leven in onze darmen, is het ook belangrijk te bepalen hoe deze bacteriën met elkaar communiceren en onze darmen beïnvloeden. Suikers blijken hierbij een belangrijke rol te spelen. Zowel de darmcellen als bacteriën zijn omgeven door structuren opgebouwd uit verschillende suikers. De combinaties van deze suikers geven specifieke complexe suikerstructuren die verschillen tussen bacteriën onderling en darmcellen: een soort van suiker streepjescode (Figuur 1).

Zowel de goede bacteriën als de slechte bacteriën zijn in staat de suiker streepjescode aan te passen doormiddel van enzymen. Deze enzymen kunnen suikers toevoegen aan de aanwezige suikerstructuren of ervan afknippen. Zo kunnen bijvoorbeeld de slechte bacteriën, zich vermommen met suikerstructuren die lijken op de menselijke suikerstructuren waardoor ze onopgemerkt blijven voor ons afweersysteem. En de goede bacteriën kunnen suikers afknippen om zichzelf te voeden (Figuur 1). Dit laat het belang zien van het bepalen van de individuele enzymen van goede en slechte bacteriën om zo bijvoorbeeld in de toekomst antibacteriële strategieën toe te kunnen passen op specifiek de enzymen van de slechte bacteriën.



Figuur 1 | Versimpelde illustratie van de suiker streepjescode op darmepitheel cellen, knippen van suiker door goede bacterie om zichzelf te voeden, en het maskeren van slechte bacteriën met menselijke suikerstructuren om het immuun systeem te omzeilen.

Het onderzoek beschreven in dit proefschrift is gericht op de enzymen die de buitenste suiker L-fucose op een structuur eraf kunnen knippen, de fucosidases. Deze fucosidases bevinden zich op de frontlinie tussen bacteriën en onze darmen. Fucosidases zijn al veelvuldig gelinkt aan ontstekingen, tumor progressie en bacteriële infecties. Onderzoekers zijn daarom bezig om manieren te ontwikkelen die ons in staat stellen fucosidases te volgen en zo hun exacte functie te achterhalen. Een van de manieren om dit te doen is door het gebruik van een bepaald type moleculair gereedschap, de zogenaamde activiteit gebaseerde probes (ABPs). In Figuur 2 is dit schematisch uitgelegd. Het binnenste van de fucosidases heeft een unieke vorm die specifiek suikerstructuren met L-fucose herkennen. Wanneer L-fucose herkent wordt door de fucosidase zal deze een chemische reactie aangaan om de L-fucose van de rest van de suikerstructuur af te knippen (Figuur 2a). Chemici gebruikten de inzichten over dit mechanisme en creëerden moleculen lijkend op L-fucose die, wanneer de fucosidase hiermee de chemische reactie aangaat, verankeren in het binnenste van de fucosidases en in staat zijn een licht signaal uit te zenden (fluorophore) (Figuur 2b). Tijdens het onderzoek beschreven in dit proefschrift zijn ABPs ontwikkeld en getest op een aantal relevante bacteriële fucosidases. Daarnaast is het effect van een fucosidase van een goede darm bacterie op een pathogene bacterie bestudeerd en zijn de suikerstructuren op een darmcel gedetailleerd in kaart gebracht.



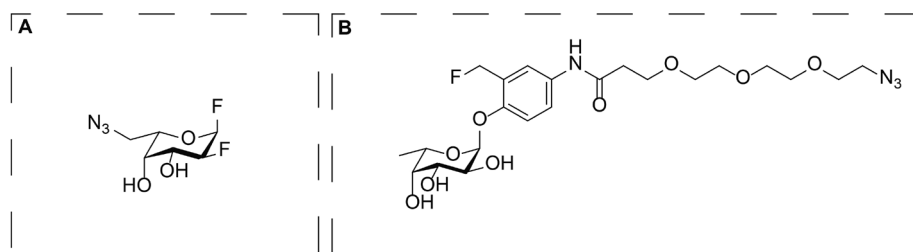
Figuur 2 | Illustratie van het fucosidase mechanisme en de werking van een ABP. a) Fucosidases herkennen suikerstructuren met L-fucose. Na enzymatisch knippen splitst de L-fucose zich af van de rest van de structuur. b) Fucosidases herkennen de op L-fucose lijkende ABP die gekoppeld is aan een fluorophore. Na de enzymatische reactie zal de ABP zich verankeren aan het fucosidase enzym.

Om het onderzoek van dit proefschrift in te leiden, wordt in **hoofdstuk 1** een gedetailleerde omschrijving gegeven van de menselijke darm en de aanwezige darmbacteriën, met een focus op L-fucose bevattende suikers en uitgescheiden fucosidases in relatie tot ziekte en gezondheid. Daarnaast bespreken we de verschillende soorten fucosidases, hun bijbehorende mechanismes en hoe dit gebruikt wordt voor het ontwerp van ABPs. In het onderzoek omschreven in **hoofdstuk 2** werd het effect van een fucosidase van een normale darm bacterie (*Bacteroides fragilis*) op de infectiviteit van een pathogene bacterie (*Campylobacter jejuni*) bestudeerd. *C. jejuni* heeft geen fucosidases om L-fucose af te knippen, maar enkele stammen (bijvoorbeeld *C. jejuni* 108) hebben wel een genetisch operon dat ze in staat stelt L-fucose op te nemen. *C. jejuni* 108 liet een toegenomen groei zien in de aanwezigheid van uitgescheiden fucosidase van de normale darm bacterie in combinatie met porcine gastric mucine (PGM, een slijm eiwit dat erg veel verschillende suikerstructuren bevat). Vervolgens werd een Caco-2 darmcel model gebruikt, dat mucine tot expressie brengt, in combinatie met een bekende fucosidase remmer om het effect van de *B. fragilis* fucosidase op *C. jejuni* 108 invasie te bepalen. *C. jejuni* 108 bleek meer invasie te laten zien in de aanwezigheid van *B. fragilis* en dit effect kon gecorreleerd worden aan fucosidase activiteit. Daaruit kon geconcludeerd worden dat deze specifieke stam *C. jejuni* zowel voor zijn groei als voor invasie afhankelijk is van de L-fucose afgeknipt door de fucosidase van de normale darm bacterie.

In het menselijk darmsysteem worden de epitheel cellen omringd door een grote hoeveelheid bacteriën. Deze bacteriën zijn in staat complexe suikerstructuren om te zetten in butyraat, welke vervolgens een belangrijke energie bron is voor de epitheel cellen. Naast de functie als energie bron blijkt butyraat ook betrokken te zijn bij de mucine expressie. Het Caco-2 cel model wordt veelal gebruikt om bacteriële interacties met mucine suikers te bestuderen. Echter, het effect van het bacteriële metabool butyraat op de suiker streepjescode *in vitro* was nog niet eerder onderzocht. Daarom is in het onderzoek dat **hoofdstuk 3** in staat gekeken naar het effect van butyraat op het suikerpatroon van de Caco-2 cellen. Door middel van een PGC nano-LC-ESI-MS/MS methode hebben we laten zien dat butyraat een specifieke suiker streepjescode geeft op de darmcellen. We hebben hiermee een basis gelegd voor een techniek die uiteindelijk gebruikt kan worden om de specifieke suikerpatronen, achtergelaten door actieve bacteriële fucosidases, in kaart te brengen.

Hoofdstuk 4 beschrijft de chemische synthese van een 2-deoxy-2-fluoro fucoside ABP (Figuur 3a). Vervolgens werd getest of dit molecuul in staat was zich te verankeren aan verschillende fucosidases. Daarnaast werd de snelheid van het verankeren aan het fucosidase onderzocht, welke overeen kwam met soortgelijke eerder gerapporteerde ABPs. De in dit onderzoek ontwikkelde efficiënte en flexibele syntheseroute maakt het in de toekomst mogelijk om makkelijk meerdere variaties te maken en daarmee snellere of specifiekere probes te creëren tegen bacteriële fucosidases.

Tijdens het onderzoek beschreven in **hoofdstuk 5** werd een ander soort fucosidase ABP gesynthetiseerd en getest. Waar de ABP uit **hoofdstuk 4** zich slechts aan één type fucosidases verankert, is deze ortho-fluoromethylphenol fucopyranoside gebaseerde probe in staat zich aan alle types fucosidases te verankeren (Figuur 3b). Verder is in deze studie onderzocht op welke specifieke posities in/op het fucosidase de probe zich verankert en wordt besproken wat de gevolgen hiervan zijn voor het gebruik van deze probe.



Figuur 3 | Structuren van de gesynthetiseerde fucosidase ABPs. a) 2-deoxy-2-fluoro fucoside ABP. b) *ortho*-fluoromethylphenol fucopyranoside ABP.

In **Hoofdstuk 6** wordt, naast de Engelse samenvatting, de mogelijke vooruitzichten van de in dit onderzoek ontwikkelde ABPs beschreven. Alsmede enkele plannen voor andere ABPs gericht op bacteriële fucosidases.

Deze Nederlandse samenvatting is gesimplificeerd omwille van de begrijpelijkheid voor de niet-chemici of niet-biologen.

Curriculum Vitae

Yvette Luijkx was born in Etten-Leur on the 21th of December 1991. She graduated from 't Raayland College in Venray in 2010 and started the bachelor Molecular Life Sciences (MLS) in Nijmegen that same year. As part of her study she conducted an internship at the department of Biomolecular chemistry within the group of Prof. G.J.M. Pruijn in 2013. During this internship she contributed to optimization of a phenylglyoxal-based visualization of citrullinated proteins on Western blot under the supervision of Dr. S. Hensen. This research resulted in a publication in *Molecules*. After obtaining her bachelors degree in 2014, she continued with the MLS master in Nijmegen, specializing in biochemistry and organic chemistry. Her first master internship was in the department of Bioorganic chemistry at the Radboud University Nijmegen in the group of Prof. van Hest under the supervision of Dr. A. Borrmann. This research resulted in the master thesis "N-terminal functionalization of the Fc fragment by Sortase A mediated reactions". Hereafter she started her second internship at the Biolink group in Stavanger (Norway) under the supervision of Prof. Einar Bakstad. She completed her master thesis "Investigation of novel highly strained molecules". After receiving her Master of Science degree in 2016, she moved to Utrecht to start her doctoral research in the Chemical Biology and Drug Discovery (CBDD) group at the Utrecht institute of Pharmaceutical Sciences (UIPS) under the supervision of Dr. T. Wennekes, Dr. K. Strijbis and Prof. G. J. Boons. The results of this work are described in this thesis. In February 2021 she started her job as a biotechnical production specialist antigens at MSD in Boxmeer.

List of publications

Bacteroides fragilis fucosidases facilitate growth and invasion of *Camylobacter jejuni* in the presence of mucins.

Luijkx, Y.M.C.A.; Bleumink, N.M.C.; Jiang, Jianbing; Overkleeft, H.S.; Wösten, M.M.S.M.; Strijbis, K. & Wennekes, T.

Cellular microbiology **2020**, <https://doi.org/10.1111/cmi.13252>

O-glycomic signatures of sodium butyrate-stimulated enterocyte differentiation

Luijkx, Y.M.C.A.; Madunic, K.; Wennekes, T.; Strijbis, K. & Wührer, M.

Manuscript in preparation

Development of a 1,2-difluorofucoside activity-based probe for profiling GH29 fucosidases

Luijkx, Y.M.C.A.; Jongkees, S.A.K.; Strijbis, K. & Wennekes, T.

Organic & Biomolecular Chemistry, **2021**, DOI: 10.1039/D1OB00054C

Detection of bacterial α -L-fucosidases with an *ortho*-quinone methide-based probe and mapping of the probe-protein adducts

Luijkx, Y.M.C.A.; Henselijn, A.J. ; Bosman, G.P.; Cramer, D.A.T.; Giesbers, C.A.P.; van 't Veld, E.M.; Boons, G.J.P.H.; Heck, A.J.R.; Reiding, K.R.; Strijbis, K. & Wennekes, T.

Manuscript in preparation

Other publications

MUC1 is a receptor for the *Salmonella* SiiE adhesion that enables apical invasion into enterocytes.

Li, X.; Bleumink-Pluym, N.M.C.; Luijkx, Y.M.C.A., Wubbolts, R.W.; van Putten, J.P.M. & Strijbis, K.

PLoS Pathogens **2019**, 15 (2), pp. e1007566

Acknowledgements

Please bear with me while I acknowledge the numerous people who helped and supported me during the duration of this PhD, without whom, I would unlikely be where I am today.

First and foremost, I would like to thank Geert Jan Boons, Tom Wennekes and Karin Stribis. Geert Jan, as promoter, thank you for the opportunity to start my PhD at the chemical biology and drug discovery group. Tom and Karin, special thanks goes out to both of you. You have been my daily supervisors for the duration of my PhD and I enjoyed to have been guided by such enthusiastic committed persons. Here I would also like to thank Jos van Putten, you were always interested in how my research was going and it was nice to know you looked out for me from the side-line.

I had the pleasure during my PhD to be welcomed in two research groups. It was nice to always have someone around to have a brainstorm with, or to just have a coffee and a chat. I enjoyed the laboutings, Christmas potluck, birthday cakes and the monthly borrels. Thanks to CBDD: Geert Jan, Margreet, Roland, John, Dirk, Ed, Seino, Tom, Arwin, Robert, Edwin, Roosmarijn, Linda, Nino, Justyna, Luca, Ivan, Apoorva, Mechman, Lifeng, Enrico, Ingrid, Hanna, Pieter, Tim, Helena, Jun Yang, Erianna, Krishna, Cyril, Xuan, Downson, Vito, Gael, Gerlof, Rosanne, Suhela, Frederik, Jack, Reshmi, Xianke, Lemeng, Javier, Yanyan, Minglong, Margherita, Diksha, Zhiyong, George, Pouya, Milka, Bernd, Cindy, Ilhan. Ofcourse also Nathaniel Martin and his group that were part of CBDD for my first years: Matthijs, Kamal, Tom, Charlotte, Emma, Nicola, Timo, Laurens. And the group that didn't hesitate to take me in as one of their own, thanks Infection biology: Jos, Marc, Karin, Linda, Nancy, Marcel, Xinyue, Celia, Xuefeng, Jiannan, Guus, Carlos, Liane, Daphne, Maytrayee, Koen. And a little bit further along the hallway, thanks to the molecular host defense group: Henk, Roel, Melanie, Albert, Hanne, Maaik, Edwin.

Special thanks to Nancy Bleumink, Linda Heijmen-van Dijk and Arwin Brouwer for their patience and steadfastness in managing the labs, keeping the labs running no matter what is thrown at you. And ofcourse for listening to me when I wanted to ask a "quick" question; which definitely weren't always quick.

Additionally without the work carried out by a number of people during my PhD, the work presented herein would be much more limited in scope, and I would like to thank all of the people that contributed towards this. To Anniek and Stefania, who carried out a large body of both synthesis work on the quinone methide and synthetic analogues of the aziridine cyclophellitol probes. To Gerlof who carried out a large body of work on expressing and purifying the fucosidases. To Dario and Katharina for the fruitful collaborations in the field of mass spectrometry. And the CCI group, Richard and Esther, thank you for the guidance and collaborations on the microscopes!

All work and no play would be quite dull, so everyone involved in organizing laser game, bowling and board games night will always have a special place in my heart: Tom, Charlotte, Rosanne, Nicola, Emma, Ingrid, Hanna and Enrico. We had several memorable events, 90's party in Tivoli, kings night at Domplein, Ingrid's wedding, and I am probably forgetting even more memorable ones. I enjoyed every one of these events and I hope we will still have many more to come.

Ingrid and Enrico, you two are not only my paranymphs, but during my PhD you have been my sounds boards and became two of my closest friends. Working in the lab was definitely better when you two were around, to discuss chemistry related issues and to lift up my mood when these great ideas and suggestions did not result in the wanted outcome.

Ingrid, when I started at CBDD you were one of the first people I met. We clicked right away and have had many Mario Kart nights, coffee breaks and parties. I love how you are always optimistic, and maybe most memorable is how thoughtful you are about the people around you. I will never forget our first symposium together on December 5th, where I woke up really nervous for my first big presentation and as a surprise you had placed the chocolate letter Y in my shoe, with a little motivating poem! I wish you all the best in your new job and all new adventures that lie ahead of you.

Enrico, we met during your first internship and got along very well, brabantse humor goes a long way. I was therefore really happy when you decided to come back as a PhD. You were my go to guy when I wanted to discuss new chemistry ideas, and a great listener when I had to blow of some steam. I enjoyed our gaming nights, Mario Kart more than Overcooked, as with Mario Kart I was actually allowed to touch a controller. You are a great chemist and an even better friend. Good luck with the last part of your PhD and all future opportunities that come to you!

And a special shout out to the people in my life outside academia: Pap (Drs. A.A.C.M. Luijkx), Mam (Sylvia), Opa Axa (Jacques), Ashley, Serena, Sharon, Quincy, Cadey, Rob, Eric, Ties, Jeroen, Luc, Kendra, Toon and Maria.

And last but definitely not least **Thanks Peter and Layla.**

