O -Mucin-degrading carbohydrate-active enzymes and their possible implication in inflammatory bowel diseases

Aurore Labourel, Jean-Luc Parrou, Céline Deraison, Muriel Mercier-Bonin,
Sophie Lajus, Gabrielle Potocki-Veronese

HAL Id: hal-04048006

https://hal.inrae.fr/hal-04048006

Submitted on 27 Mar 2023

PORTLAND PRESS

Review Article

O-Mucin-degrading carbohydrate-active enzymes and their possible implication in inflammatory bowel diseases

Aurore Labourel[1], Jean-Luc Parrou[1], Céline Deraison[2], Muriel Mercier-Bonin[3], Sophie Lajus[1] and
Gabrielle Potocki-Veronese[1]

[1]TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France; [2]IRSD, Université de Toulouse, INSERM, INRAE, ENVT, Univ Toulouse III - Paul Sabatier (UPS), Toulouse, France; [3]Toxalim, Research Center in Food Toxicology, INRAE, ENVT, INP-Purpan, UPS, Université de Toulouse, 31000, Toulouse, France

Correspondence: Gabrielle Potocki-Veronese (veronese@insa-toulouse.fr)

Inflammatory bowel diseases (IBD) are modern diseases, with incidence rising around the world. They are associated with perturbation of the intestinal microbiota, and with alteration and crossing of the mucus barrier by the commensal bacteria that feed on it. In the process of mucus catabolism and invasion by gut bacteria, carbohydrate-active enzymes (CAZymes) play a critical role since mucus is mainly made up by *O*- and *N*-glycans. Moreover, the occurrence of IBD seems to be associated with low-fiber diets. Conversely, supplementation with oligosaccharides, such as human milk oligosaccharides (HMOs), which are structurally similar to intestinal mucins and could thus compete with them towards bacterial mucus-degrading CAZymes, has been suggested to prevent inflammation. In this mini-review, we will establish the current state of knowledge regarding the identification and characterization of mucus-degrading enzymes from both cultured and uncultured species of gut commensals and enteropathogens, with a particular focus on the present technological opportunities available to further the discovery of mucus-degrading CAZymes within the entire gut microbiome, by coupling microfluidics with metagenomics and culturomics. Finally, we will discuss the challenges to overcome to better assess how CAZymes targeting specific functional oligosaccharides could be involved in the modulation of the mucus-driven cross-talk between gut bacteria and their host in the context of IBD.

Introduction

Intestinal mucus is notably formed by glycoproteins, the *O*- and *N*-linked glycans which make up to 80% of their total mass. These structurally complex glycans influence the cohesiveness of the mucus network. Their structure and defects in their biosynthesis [1] may be differently modulated by environmental factors [2–4], in particular in inflammatory bowel diseases (IBD), such as Crohn's disease (CD) and ulcerative colitis (UC). These diseases, which are more prevalent in the modern world, are associated with the Western diet with low gut microbiota-accessible carbohydrates, which is linked to a lack of microbial diversity and altered production of short-chain fatty acids [5]. IBD have become increasingly prevalent in Europe, North America and Australia since the 1950s, with their incidence rising beyond the western world over the past 30 years [6]. They are characterized by an imbalance in the composition of the intestinal microbiota [7–9], the so-called 'dysbiosis', and by an uncontrolled inflammatory response to luminal content [10]. In addition, the intestinal barrier function is severely impaired: significant changes in mucus secretion and mucin *O*-glycosylation

occur, the inner mucus barrier is thinner and looser, and the space closest to the epithelium is invaded by commensal bacteria and/or enteropathogens [11–14]. Indeed, together with the dietary glycans (HMOs for infants, and then fibers of plant origin) and, to a minor extent, with microbial exopolysaccharides, the mucus glycans are the main carbon sources for gut bacteria. To deal with the huge structural diversity of glycans they feed on, and to break them down into metabolizable monosaccharides, intestinal bacteria produce a large panel of carbohydrate-active enzymes (CAZymes) with various specificities [15], and a battery of proteins to sense, bind and transport glycans into bacterial cells for their complete breakdown. These metabolic machineries are often encoded by multigenic clusters, defined as polysaccharide utilization loci (PUL) in the Gram-negative Bacteroidota (formerly Bacteroidetes [16]), one of the dominant phyla in the human gut. Bacteroidota include bacteria that are able to feed on mucins, such as the commensal species *Bacteroides thetaiotamicron* and *Bacteroides fragilis*, the abundance of which seems to be correlated to CD and UC [17,18]. CAZymes targeting mucins are also found in other Gram-negative mucin degraders, such as, for example, *Akkermansia muciniphila* [19], which is negatively associated with IBD [8], and in Gram-positive ones, such as *Ruminococcus gnavus*, which is, on the contrary, positively associated with these diseases, and CD in particular [20]. Given that mucin degraders can be both positively and negatively correlated to IBD, it is very difficult to determine whether mucin breakdown by gut bacterial CAZymes is involved in intestinal epithelium inflammation. CAZymes produced by mucin-degrading bacteria have been reviewed several times over the last decade [21–25]. Nevertheless, some recent functional metagenomics studies have highlighted the potential of CAZymes produced by uncultured gut bacteria, and appearing as IBD biomarkers, to degrade human *O*-glycans present in mucus [26,27] and in human milk oligosaccharides (HMOs). These oligosaccharides could indeed compete with the *O*-glycans of mucins lining the intestinal epithelium towards mucus-degrading CAZymes, since they are structurally similar to their extremities. HMOs seem to protect against IBD [28–30], although these data are still controversial [31], as for other functional foods that are metabolized thanks to CAZymes from the intestinal microbiome [32].

In this review, we will update the list of CAZymes that have been shown to target *O*-mucins, since mucins are much more sparsely *N*-glycosylated than *O*-glycosylated [33,34]. Here, we will discuss the advantages and limits of the methods that are currently available to identify them within the entire microbiome. Finally, we will explore the potential role that mucin-degrading CAZymes could play in the interrelationships between the host, diet and the microbiota in the context of gut inflammation.

Mucin-degrading CAZymes
Structure of the O-mucins
MUC2, the major colonic mucin is composed of ∼20% protein and ∼80% glycan. The protein sequence rich in Ser/Thr residues are primary targets for *O*-glycosylation with *N*-acetyl-D-galactosamine (GalNAc), creating the foundation upon which long oligosaccharide chains are built. While these side chains are composed of only five different monosaccharides (*N*-acetyl-D-glucosamine (GlcNac), GalNac, D-galactose, L-fucose and *N*-acetylneuraminic acid [also called sialic acid, Neu5Ac]), the order in which they can be assembled is hugely variable. In addition, the GlcNac and galactose residues can be sulfated, adding a supplemental level of complexity. The resulting *O*-glycans extend out from the mucin protein core in a 'bottlebrush' configuration and its heterogeneity provides some resistance to microbial degradation [33,35]. It should be emphasized that the mucin *O*-glycans vary enormously along the intestinal tract and between species (Table 1), compelling researchers in the area to take greater care when studying bacteria and mucin from the same niche.

Identification and characterization of O-mucin degrading CAZymes
Several comparative genomic analyses have attempted to determine the CAZyme mucin-degrading profiles of the human gut microbiota [37,38], and there have been efforts to couple these data with transcriptomic, proteomic, genetic and biochemical studies in order to understand mucin catabolic pathways. One of the earliest and most studied commensal bacteria is *B. thetaiotaomicron*, which is one of the richest human bacteria in glycan-degrading enzymes [15,39]. In 2008, Martens and his team performed whole-genome transcriptional profiling and genetic tests to define the mechanisms underlying host glycan foraging *in vivo* (mice) and *in vitro* [40]. Mucin *O*-glycans are the principal substrate foraged *in vivo* and 16 PULs were identified with probable specificities for *O*-glycans. Recently, two of these PULs were characterized highlighting the major role of GH16 enzymes [41] and sulfatases [42–44]. Growth assays on mucins combined with transcriptomic analyses were also performed on other Bacteroidota such as *B. fragilis* [45,46] and *B. massiliensis* [45]. Similar studies were performed on the mucin degraders *A. muciniphila* [47–50] (Verrucomicrobia), *R. gnavus* [51] (Firmicutes) and *Bifidobacterium bifidum* [52] (Actinobacteria). This list is destined

Table 1 Main characteristics of substrates

Substrates categories	Representative structures
Chromogenic /fluorogenic glycoside	
Origin: synthetic substrates	X-GlcNAc, X-sulfate, X-Neu5Ac, …
Commercial price: a few hundreds to a few thousand €/g	*p*NP-Fuc, *p*NP-GlcNAc, …
Applications: conventional screening; specificity characterization in microplates; droplet Microfluidics-based screening (fluorogenic substrates)	Fluorescein-GlcNAc, …
HMOs and blood antigens	
Origin: synthetic substrates	
Commercial price: tens of thousands €/g	
Applications: specificity characterization (HPAEC-PAD); droplet microfluidics-based screening and culturomics	
Pig gastric mucin (PGM)	
Origin: sampling on animals	
Commercial price: €1 to €10/g	
Applications: transcriptomics; characterization of synthetic cocktails of CAZmes	
Human colonic mucin	
Origin: surgical pieces, cell lines, organoids	
Not commercial	
Applications: ELISA assays; characterization of synthetic cocktails of CAZymes (though not exemplified yet)	

Symbols: GlcNac ■ (blue), Gal ● (yellow), GalNAc ■ (yellow), Fucose ▲ (red), Sialic acid ◆ (purple), Glc ● (blue), Sulfate **S**; (a) Representative, not exhaustive structures decorating PGM chains (mainly 4–6 units), leading to extreme diversity. (b) Schematic view of the most common, core 3-based branch (adapted from [36]). For the sake of clarity, the nature of the bonds between the units are not depicted.

to get longer as 23 representative gut microbes were shown to grow on porcine intestinal mucin as their sole carbon source [38].

The *O*-mucin-degrading CAZymes identified by genomic and transcriptomics analysis are usually biochemically characterized using porcine gastric and intestinal mucins, bovine submaxillary mucin, mouse crude mucin as well as HMOs, blood-group oligosaccharides and synthetic oligosaccharides. To date, 18 glycoside hydrolase (GH) families were shown to be involved in the degradation of *O*-mucin glycans (www.cazy.org). GH33 sialidases cleave terminal sialic acid residues while fucose residues are removed by GH29 or GH95. Subsequent degradation of the glycan chains involves lacto-*N*-biosidase (GH136), lacto-*N*-biose phosphorylases (GH112), *N*-acetyl-glucosaminidases (GH84, GH85, GH89, GH20), *N*-acetyl-galactosaminidases (GH31, GH101, GH109, GH129) and galactosidases (GH2, GH35, GH42, GH98). There are also endo-acting *O*-glycanases (GH16), which can cleave large glycan structures. Table 2 lists the characterized *O*-mucin-degrading CAZymes since 2015. It is an update of a previous review [22]. To provide an overview of the complete catabolic pathway of *O*-mucins, the proteases that have been shown to be specific to this substrate are listed in Table 3, and the sulfatases and the sialate *O*-acetylesterase are listed in Table 4.

Overall, about a hundred mucin-degrading CAZymes have been biochemically characterized with substrates of defined structure, allowing the determination of their specificity towards the host and dietary glycans that could compete to feed gut bacteria. This limits the understanding, at the molecular level, of the relationships between the microbiota, its host and diet. In addition, these biochemical data are far from being exhaustive, especially when we consider that uncultured species dominate in the human gut microbiota [75]. Recently, activity-based metagenomics [76] have enabled the capture of the first CAZymes from uncultured gut bacteria (α-fucosidases, β-*N*-acetyl-galactosidases, β-galactosidases, β-*N*-acetyl-glucosidases, β-*N*-acetyl-neuraminidases, α- and β-mannosidases), which have proved to affect the structure of various *O*-, *N*-glycans, HMOs and human colonic mucus [26,27,77]. These studies also provided the first evidence of the over-representation of the genes coding for some of these enzymes in the human gut metagenome of patients suffering from IBD.

It is noteworthy that metagenome assemblies generated from the gut microbiota include a wide range of other organisms, apart from bacteria, such as archaea, eukaryotes and viruses that warrant a more thorough investigation

Table 2 The list of biochemically characterized *O*-mucin degrading CAZymes

CAZy family	Name or locus tag	Activity	Organism	Accession numbers
GH2	Amuc_0824	β-galactosidase (preference for galactose-β-1,3-*N*-acetyl-D-galactosamine)	*Akkermansia muciniphila* ATCC BAA-835	ACD04657.1 [53]
	Amuc_1666	β-galactosidase (preference for galactose-β-1,4-*N*-acetyl-D-glucosamine)		ACD05484.1 [53]
GH16	BT2824	*O*-Glycan β-1,4-galactosidase with a requirement for a β1,3-linked sugar at the -2 position	*Bacteroides thetaiotaomicron* VPI-5482	AAO77930.1 [41]
	BF4058		*Bacteroides fragilis* NCTC 9343	CAH09732.1 [41]
	BF4060			CAH09734.1 [41] PDB 6T2S
	BACCAC_02679; CGC64_00695		*Bacteroides caccae* ATCC 43185	ASM64646.1 [41]
	BACCAC_02680; CGC64_00690			ASM64645.1 [41] PDB 6T2Q, 6T2P, 2T2O
	BACCAC_03717; CGC64_17650			ASM67518.1 [41] PDB 6T2R
	Amuc_0724		*Akkermansia muciniphila* ATCC BAA-835	ACD04561.1 [41] PDB 6T2N
	Amuc_875			ACD04708.1 [41]
	Amuc_2108			ACD05917.1 [41]
GH29	*Ri*Fuc29, ROSEINA2194_01891/01890	α-1,4-fucosidase	*Roseburia inulinivorans* DSM 16841	C0FT24/C0FT23 (Uniprot) [54]
	E1_10125; RUGNEv3_10125	α-L-fucosidase (preference for α1,3/4 fucosylated linkages)	*Ruminococcus gnavus* E1	A0A6N3BKT0 [55] PDB 6TR3, 6TR4
	E1_10180; RUGNEv3_10180			- [55]
	ATCC_03833; RUMGNA_03833	α-L-fucosidase (preference for α1,2 linkages)	*Ruminococcus gnavus* ATCC 29149	- [55]
	Afc2; CPF_2130	α-L-fucosidase (preference for α1,3/4 fucosylated linkages)	*Clostridium perfringens* ATCC 13124	ABG83106.1 [56]
	*Am*GH29A, Amuc_0010	α-L-fucosidase (preference for Fucα1,3GlcNAc)	*Akkermansia muciniphila* ATCC BAA-835	ACD03857.1 [57]
	*Am*GH29B, Amuc_0146			ACD03990.1 [57]
	*Am*GH29C, Amuc_0392	α-L-fucosidase (preference for α1,3/4 fucosylated linkages)		ACD04231.1 [57]
	*Am*GH29D, Amuc_0846			ACD04679.1 [58]
GH31	*Bp*GH31	α-*N*-acetylgalactosaminidase	*Phocaeicola plebeius* (formerly *Bacteroides plebeius*)	WP_007559952.1 [58]
	*Bc*GH31		*Bacteroides caccae*	WP_005682123.1 [58]
GH33	*Rg*NanH	Intramolecular *trans*-sialidase (IT-sialidase) producing 2,7-anhydro-Neu5Ac selectively from α2-3-linked sialic acid substrates	*Ruminococcus gnavus* ATCC 29149	ETD19277.1 [59]
	NanH, BF1806	Sialidase	*Bacteroides fragilis* NCTC 9343	WP_010992682.1 [60]
	*Am*GH33A, Amuc_0625	α2,3- and α2,6-sialyl linkages	*Akkermansia muciniphila* ATCC BAA-835	ACD04462.1 [57]
	*Am*GH33B, Amuc_1835			ACD05653.1 [57]
GH35	*Am*0874; Amuc_0771	β-galactosidase (β1,3- and β1,6-linked galactose residues from various substrates, preference for galactose-β-1,3-*N*-acetyl-D-glucosamine)	*Akkermansia muciniphila* ATCC BAA-835	ACD04606.1 [53,61]
	Amuc_1686	β-galactosidase (Galβ1-3GalNAc)		ACD05504.1 [62]
GH95	*Ri*Fuc95, RO-SEINA2194_01889/01888/01887	α-1,2-fucosidase	*Roseburia inulinivorans* DSM 16841	C0FT22/C0FT21/C0FT2 (Uniprot) [54]
	ATCC_00842; RUMGNA_00842	α-L-fucosidase (preference for α1,2 linkages)	*Ruminococcus gnavus* ATCC 29149	[55]
	E1_10587; RUGNEv3_10587	α-L-fucosidase (low activity, not possible to assess its substrate specificity)	*Ruminococcus gnavus* E1	[55]
	Afc3; CPF_2129	α-L-fucosidase (preference for α1,2 linkages)	*Clostridium perfringens* ATCC 13124	ABG82552.1 [5656]
	*Am*GH95A, Amuc_0186	α-L-fucosidase (preference for α1,2/3 fucosylated linkages)	*Akkermansia muciniphila* ATCC BAA-835	ACD04030.1 [57]
	*Am*GH95B, Amuc_1120			ACD04946.1 [57]

Continued over

Table 2 The list of biochemically characterized *O*-mucin degrading CAZymes (Continued)

CAZy family	Name or locus tag	Activity	Organism	Accession numbers
GH98	*Ri*GH98, ROSEINA2194_02198	Blood-group endo-β-1,4-galactosidase	*Roseburia inulinivorans* DSM 16841	C0FTX7 (Uniprot) [54]
GH109	*Bv*GH109_1	α-*N*-acetylgalactosaminidase	*Bacteroides vulgatus*	WP_005848863.1 [58]
GH112	*Rh*GLnbp112, RHOM_04120	β-1,3-galactosyl-*N*-acetylhexosamine phosphorylase	*Roseburia hominis* A2-183	AEN95946.1 [54]
	*Ri*GLnbp112, ROSEINA2194_01885		*Roseburia inulinivorans* DSM 16841	EEG94248.1 [54]
GH136	*Rh*Lnb136I, RHOM_04110 + *Rh*Lnb136II, RHOM_04115	Lacto-*N*-biosidase	*Roseburia hominis* A2-183	G2T0V0 + G2T0V1 (Uniprot) [54]
	*Ri*Le$^{a/b}$I136, ROSEINA2194_01899 + *Ri*Le$^{a/b}$II136, ROSEINA2194_01898	Lacto-*N*-biosidase. Releases Lewis a triose or Lewis b tetraose from fucosylated HMOs including lacto-*N*-fucopentaose II, lacto-*N*-difucohexaose I and lacto-*N*-difucohexaose II	*Roseburia inulinivorans* DSM 16841	C0FT31 + C0FT32 (Uniprot) [54]
	*Er*Lnb136, HMPREF0373_02965	Lacto-*N*-biosidase	*Eubacterium ramulus* ATCC 29099	ERK42296.1 [54] PDB 6KQS, 6KQT
New family to be assigned	Amuc_1547	Exclusive specificity towards the sialyl-T-antigen but accepts α2,6-sialylation or β1,6-substitution of the GalNAc unit	*Akkermansia muciniphila* ATCC BAA-835	ACD05368.1 [57]

[78]. For instance, gut commensal fungi are likely involved in IBD [79] but to our knowledge, none of their CAZymes has been characterized to date.

Technological bottlenecks and challenges in the discovery and characterization of mucin-degrading CAZymes

The combination of growth assays on mucins combined with transcriptomics, or proteomics [80] is currently the best strategy to highlight the CAZymes involved in mucin degradation by targeted species. These species tend to be either highly prevalent commensals or species that are associated with intestinal diseases [51].

In addition, this approach can be used to identify the enzymes from bacterial mini-consortia that act synergistically to break down the mucin network, or, on the contrary, that compete to forage on mucin glycans [81]. It is nevertheless restricted to the study of cultured species, and although 'artificial' consortia could be studied, they do not represent the diversity and the complexity of the bacterial and enzymatic machineries that are involved in mucus breakdown in the gut microbiome. Although those that exist are extremely interesting, there are currently very few metatranscriptomic and/or metaproteomic studies available [82,83] giving access to the battery of CAZyme-encoding genes expressed during mucin consumption by the entire microbiota, in particular in an IBD context [84].

Technically speaking, performing growth assays and transcriptomics or proteomics studies with biological and technical replicates requires grams of substrate, which is easy when such studies are performed on widely available natural substrates such as plant cell wall polysaccharides. When the focus of the studies is human mucin degradation, the challenge is very different because of the impossibility of sampling human intestinal mucus in large amounts. The solution currently used is to perform growth studies on porcine mucins, which share structural similarities with human mucin core proteins. Unfortunately, however, there are important differences in the relative abundance of core types, neutral, acidic or sulfated glycans [85] (Table 1), which are crucial in terms of enzymatic accessibility to the network of mucins. In brief, the main structural characteristics that differentiate pig gastric mucin (PGM) – one of the mostly used substrates for *in vitro* studies – from human colonic mucin are respectively [36,86–88]: MUC5 dominant versus MUC2 enriched mucin; cores 2 and 1 versus core 3 dominant; low degree of sialylation versus extensive sialylation (mono, di tri) and sulfate groups, leading to acidic human colonic mucus; extensive presence of blood type ABH and Lewis epitopes and i/I branches versus extensive presence of the Sda/Cad antigen (GalNAc(NeuAc-)Gal) and low abundance, mono-fucosylation, including blood group ABH antigens; galactose as the most common terminal residue, as well as presence of a sulfate group, approximately 50% of termini for both in PGM. In addition, using such complex polymeric structures does not allow for identification of the specificity of the expressed CAZymes towards a particular oligosaccharidic motif or linkage type, except using powerful but complex techniques combining

Table 3 The list of proteases that have been shown to be involved in *O*-mucin degradation

Protease family	Name or locus tag	Activity	Organism	Accession numbers
Serine protease Pic (Protein involved in Intestinal Colonization) autotransporter	Pic	Active on bovine submaxillary mucin and mouse crude mucin. Similar cleavage pattern to BT4244 but no tolerance for the presence of sialic acid	*Escherichia coli* (EAEC) 042	AF097644.1 [63]
	Pic	Active on bovine submaxillary mucin and mouse crude mucin	*Shigella flexneri* 2457T	Q54151 (Uniprot) [63]
M60-like zinc metalloprotease	BT4244	Cleaves peptides N-terminally to glycosylated Ser or Thr residues and preferred the modified residue to bear atruncated, tumor-associated glycan such as the Tn (*O*-GalNAc) or T (*O*-GalNAc-galactose) antigen. Sensitivity to sialic acid	*Bacteroides thetaiotaomicron* VPI-5482	WP_008764444.1 [64–66] PDB 5KD2, 5KD5, 5KD8
	SslE	Active on bovine submaxillary mucin	*Escherichia coli* IHE3034	WP_001034562.1 [67]
	YghJ	Active on human MUC2 and MUC3	*Escherichia coli* H10407 ETEC	E3PJ90 [68] (Uniprot)
	AM0627	Cleaves between two *O*-glycosylated Ser or Thr residues bearing atruncated, tumor-associated glycan such as the Tn (*O*-GalNAc) or T (*O*-GalNAc-galactose) antigen. Sensitivity to sialic acid	*Akkermansia muciniphila* ATCC BAA-835D-5	B2UPI7 (Uniprot) [42,69,70] PDB 7SCI, 7YX8
	AM0908	Cleaves N-terminally to a glycosylated Ser or Thr with a marked preference for an adjacent Ser or Thr residue. Tolerate sialic acid		B2UQK5 (Uniprot) [66]
	AM1514	Cleaves N-terminally to a glycosylated Ser or Thr with a marked preference for an adjacent Ser or Thr residue. Specific for *O*-GalNAc-galactose and *O*-GalNAc antigen-modified amino acids.		B2ULE8 (Uniprot) [66]
	PA0572, IMPa	Hydrolyze a minimally *O*-glycosylated substrate with only the Tn- or TAg-glycans. Highest affinity when this core modification is extended with a β-1,6-linked GlcNAc (GlcNAcβ1-6GalNAcα1-Thr).	*Pseudomonas aeruginosa* (PAO1)	Q9I5W4 (Uniprot) [65,71] PDB 5KDV, 5KDW, 5KDX
	ZmpB, CPF_1489	Absolute requirement for the α-2,6-Neu5Ac present on the peptide-linked GalNAc.	*Clostridium perfringens* ATCC 13124	A0A0H2YN38 (Uniprot) [65,72] PDB 5KDJ, 5KDN, 5KDS, 5KDU
	ZmpC, CP4_3468	Action on similar protein than ZmpB 1 (75% sequence identity)	*Clostridium perfringens* CP4	F8UNJ8 (Uniprot) [72] PDB 6XSZ, 6XT1
M66-like zinc metalloprotease	StcE	Reduce the inner mucus layer in human colonic mucosal biopsies and the MUC2 glycoprotein levels in mucin-producing LS174T colon carcinoma cells	*Escherichia coli* (EHEC)	O82882 (Uniprot) [73]
Peptidase_M43 domain-containing protein	OgpA, Amuc_1119, OpeRATOR® and GlycOCATCH®	Active on Fetuin and IgA. Exclusively hydrolyzes the peptide bond N-terminal to serine or threonine residues substituted with an *O*-glycan	*Akkermansia muciniphila* ATCC BAA-835	ACD04945.1 [74] PDB 6Z2D, 6Z2O

high-performance anion-exchange chromatography with pulsed amperometric detection (HPAE-PAD), mass spectrometry and liquid chromatography-mass spectrometry (LC-MS) to analyze the structure of the oligosaccharides that are released during bacterial consumption of mucins or their *in vitro* enzymatic hydrolysis [41,89].

Table 4 The list of sulfatases and sialate *O*-acetylesterase that have been shown to be involved in *O*-mucin degradation

| Sulfatase family	Name or locus tag	Activity	Organism	Accession numbers
Sulfatase S1_4	BT3487	6S-Gal	*Bacteroides thetaiotaomicron* VPI-5482	AAO78593.1 [43]
	BT4683	3S-Gal		AAO79788.1 [43] PDB 7ALL
Sulfatase S1_11	BT1628	6S-GlcNAc	*Bacteroides thetaiotaomicron* VPI-5482	AAO76735.1 [43]
	BT3177			AAO78283.1 [43,44] PDB 7P24
Sulfatase S1_15	BT1624	6S-Gal/GalNAc	*Bacteroides thetaiotaomicron* VPI-5482	AAO76731.1 [43,44] PDB 7OZE
	BT3109			AAO78215.1 [43,44] PDB 7OZC
	BT4631			AAO79736.1 [43,44] PDB 7P26
Sulfatase S1_16	BT3057	4S-Gal/4S-GalNAc	*Bacteroides thetaiotaomicron* VPI-5482	AAO78163.1 [43,44] PDB 7OZ9
	BT3796			AAO78901.1 [43,44] PDB 7OZA
Sulfatase S1_20	BT1622	3S-Gal preferentially binding *N*-acetyl-D-galactosamine (GalNAc) and cleaving 3S-GalNAc	*Bacteroides thetaiotaomicron* VPI-5482	AAO76729.1 [43] PDB 7ANB, 7ANA
	BT1636	3S-Gal required enzyme to access colonic mucin by *Bth*		AAO76743.1 [43] PDB 7AN1, 7OQD
Sulfatase S1_46	BT1918	3S-GlcNAc, using the *N*-acetyl group as an absolute specificity determinant	*Bacteroides thetaiotaomicron* VPI-5482	AAO77025.1 [43,44] PDB 7OZ8
Sialate *O*-acetylesterase	EstA	Active on 9-*O*-acetylated sialoglycans	*Bacteroides fragilis* NCTC 9343	WP_005794991.1 [60]

Screening and characterizing the substrate specificity of CAZymes can be easily performed using chromogenic glycosides mimicking the mucin chain extremities, such as *p*NP- or X-sugars. However, there are some mucin-acting CAZymes that do not act on these artificial substrates, leading to false negative results. It is the case, for example, of the AfcA GH95 α-fucosidase from *Bifidobacterium longum* JCM1254, which is inactive on *p*NP-fucose [90]. Other GH95 and GH29 α-fucosidases act more or less efficiently on this substrate, with catalytic efficiencies varying by up to four orders of magnitude between GH29 enzymes [55].

On the contrary, the rate of false positives might be high in screening campaigns using unpurified enzymes and chemically modified substrates, that can be recognized by the native glycosidases produced by the recombinant host. For example, when screening *Escherichia coli* recombinant libraries for β-galactosidase activity, the LacZ GH2 β-galactosidase is highly problematic. *E. coli* also produces other glycosidases, such as the GH3 AWY88947 β-*N*-acetylhexosaminidase, resulting in background activity that has to be systematically taken into account to get information on the ability, or not, of the target enzymes to hydrolyze the substrates under study [27].

This is of particular importance for screening genomic or metagenomic libraries for activities of mucin hydrolysis, using conventional solid-plates or micro-plate screens. Using a battery of chromogenic glycosides in primary screens to represent the structural diversity of mucin side chains, our group discovered several mucin-targeting PUL-like multigenic clusters from a metagenomic library constructed from the mucosal-associated ileal microbiota [26]. These metagenomic PULs, cloned in fosmids and expressed in *E. coli*, encode a cocktail of synergistic CAZymes that were proven, by ELISA assays using fluorescein-conjugated lectins, to affect the structure of human colon mucus sampled from surgical pieces. In order to circumvent the use of clinical samples, validation screening assays and substrate specificity profiling could also be performed using either mucus-secreting human cell lines [41,91], cell-based platforms displaying tunable structures and patterns of *O*-glycans [92], or even – although this has not been performed to date – organoids on a 2D culture system [93,94].

Regarding the rapid identification of mucin-degrading CAZymes, metagenomics has recently been paired with microfluidics, and a fluorogenic glycoside substrate was used in a novel droplet microfluidics workflow developed to directly access metagenomic PULs cloned in fosmids and expressed in *E. coli* [27]. Indeed, droplet microfluidics screening makes it possible to perform 10^6 assays per hour, with <1 mg of substrates, which is 1000 times faster and
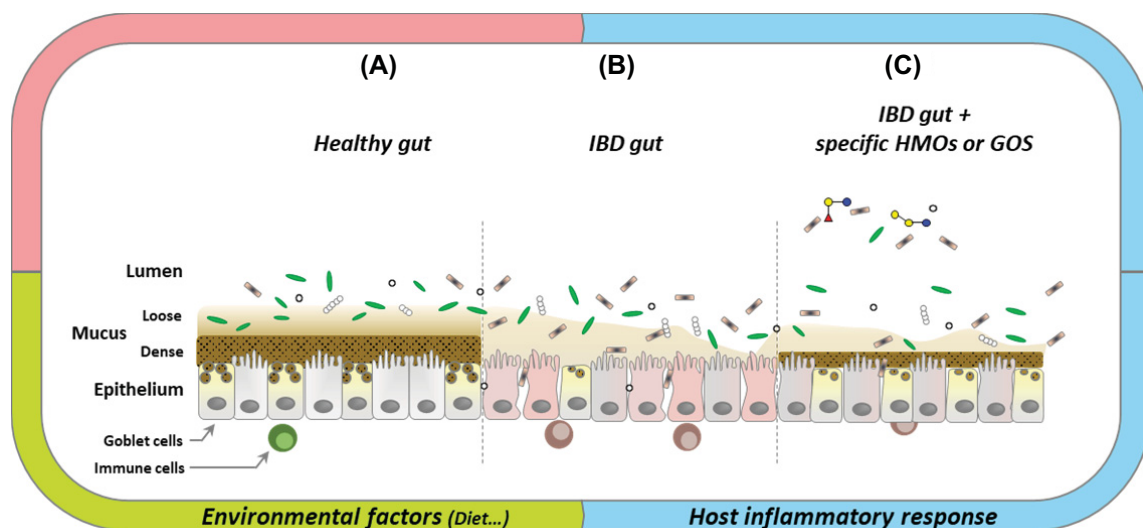
Figure 1. Schematic representation of bowel intestinal epithelium and lumen in healthy (A), IBD condition (B) or IBD condition in presence of HMOs or GOS (C)

In the healthy condition (**A**), both mucus layers (in dark, inner and sterile layer and light brown, outer layer) are present. Immune cells are present but not activated. In this condition, mucus secretion and mucus degradation by microbiota are balanced. In the IBD condition (**B**), mucus thickness is highly reduced and both mucus layers are invaded by microbiota, epithelial cell functions are impaired with loss of tight junctions. Immune cells are activated and dysbiosis is observed. Bacteria possibly infiltrate the epithelium. The origin of this unbalanced state could be defects in mucin *O*-glycosylation, resulting in lower density, higher penetrability and degradability of the inner mucus gel, the secretion of which is also affected [12,14]. With a diet supplemented with Human Milk Oligosaccharides (HMOs) and/or Galacto-Oligosaccharides (GMOs) in IBD condition (**C**), an improvement of the situation might be observed. Key: green bars and gray circles: commensal bacteria; dark bars: pathogenic bacteria.

less expensive than conventional technologies. This technology is thus compatible with the fast exploration, at low cost, of large sequence spaces to identify the target enzymes. It is based on the detection and sorting of droplet hits in function of their fluorescent or absorbance levels [27,95], bearing in mind that absorbance detection is less sensitive than fluorescence detection by orders of magnitude.

Using such workflows requiring a chemically modified substrate to identify the droplet where the target enzymatic reaction occurs, validation of activity on non-chemically modified substrates (such as commercial oligosaccharides with defined structures, including HMOs and blood antigens, that are both mucin-like motifs) is crucial. In any case, profiling substrate specificity using such structurally defined oligosaccharides, is highly expensive. In primary screening, it is thus very tempting to use such mucin-like oligosaccharides with defined structures, which represent the structural complexity and diversity of physiological substrates. It would prevent screening bias and the costly and time-consuming validation steps on real substrates. Of course, screening large libraries with such oligosaccharides using solid-plates or micro-plate screens is not conceivable, given the price of these substrates. However, recently some ultra-miniaturized droplet-based microfluidics workflows have been developed, based on positive selection of bacteria (recombinant [96] or native bacteria [97]) on native, non-chemically modified oligosaccharides. Until now, these technologies have only been exemplified for the activities of metabolization of plant-derived dietary fibers. What is more, they have not so far led to the discovery of novel functional genes, since the proofs of concepts have been established with a mini-metagenomic library, or with a microbiota sample, for which the taxonomical marker genes have only been sequenced after selecting fiber-metabolizing bacteria. Nevertheless, the development of these workflows opens some very interesting prospects for ultra-high throughput functional genomics, metagenomics and culturomics studies (using single cells or bacterial consortia), targeting the activities of mucin degradation.

Microbiota-host-diet cross-talk in IBD: the intestinal mucus layer as a key feature

Several approaches have been used over the last decade to study the links between diet, gut bacteria and their enzymes, mucus structure/penetrability, and IBD [98]. In particular, key relationships have been established between the protease activities produced by both intestinal epithelium and opportunistic pathogens, mucosal biofilm structure, and IBD.

Regarding CAZymes from the intestinal microbiome, pioneer transcriptomics studies targeting *B. thetaiotamicron* (*Bth*) revealed the potential role of these enzymes in gut inflammation, in relationship with diet. Indeed, mucin-targeting *Bth* PULs were shown to be up-regulated in fiber-deprived diets, leading to an alteration of the mucus layer, which promoted the implantation of pathogens and/or commensal opportunists [5,99]. These PULs were also shown to be up-regulated with some particular prebiotics, such as β-galacto-oligosaccharides (GOS), and HMOs [80,100]. GOS are plant-derived oligosaccharides, but as *O*-mucins, they contain β-linked galactosyl units that could be recognized and hydrolyzed by some β-galactosidases or promiscous β-hexosidases or -hexosaminidases of the gut microbiome. GOS have been tested as prebiotics on experimental colitis in rats, but no impact on inflammation was observed [101]. Moreover, the effects of other prebiotics containing β-linked galactosyl units, such as lactulose, on bacterial translocation and inflammation are still unclear [32]. In contrast, consumption of HMOs, which are much more similar to mucins than GOS, seems to protect against IBD [28–30]. Due to their structural convergence, HMOs are indeed thought to act as excellent competitor substrates of mucins towards mucin-degrading CAZymes [102] (Figure 1). In addition, certain HMOs, such as 2′-fucosyllactose, can competitively inhibit the binding of enteric pathogens to epithelial cell membranes, thus modulating specific pro-inflammatory signaling molecules [30]. In healthy adults, 2′-fucosyllactose is well tolerated and promotes the growth of beneficial bifidobacteria [103]. Furthermore, infants fed with a formula containing 2′-fucosyllactose have lower inflammatory cytokine levels. In contrast, other HMOs, such as 3′-sialyllactose, have been shown to directly interact with the host, resulting in structural modification of glycans, or even in pro-inflammatory stimulation [31]. In conclusion, with such a diversity of structures and mechanisms of interaction with both the gut microbiota and the host [104], the effect of functional oligosaccharides like GOS prebiotics and HMOs in the triggering or in the prevention of inflammation is today unclear, and too few CAZymes targeting these motifs in the intestinal microbiome have been characterized to date.

Summary

- There is a clear correlation between IBD, alteration of the colonic mucus and its barrier effect, and dysbiosis of the intestinal microbiota.

- To date, a few hundred mucus-degrading CAZymes from about ten prevalent gut bacterial species have been biochemically characterized, and only a handful of CAZymes active on mucus or mucin-motifs have been identified from uncultured bacteria, demonstrating the lack of knowledge about mucin catabolic pathways in the gut microbiome.

- The difficulties of mucus sampling, mucus interspecies structural diversity, and the cost and availability of oligosaccharides that properly mimic mucin structure, limit the identification and biochemical characterization of mucin-degrading CAZymes, and of their synergistic action to break down the mucus network.

- Novel droplet microfluidics-based screening technologies developed recently are a great opportunity to explore, at ultra-high throughput and at low cost, the mucin-degrading functions of the gut microbiome, using either functional genomics, metagenomics or culturomics.

- Some genes encoding mucin-degrading CAZymes are more prevalent in the microbiome of IBD patients compared with that of healthy individuals, although the role played by these biomarkers in inflammation has never been studied.

- The action of mucin-targeting CAZymes on mucus invasion by gut bacteria and on the host response should be explored, using a combination of *in vitro*, *ex vivo* and *in vivo* approaches specifically designed to analyze mucus characteristics and inflammation. Also, the protective role of functional foods, more specifically that of prebiotics and HMOs competing with mucin glycans towards mucus-degrading CAZymes, should be assessed. Such holobiomics studies should help demonstrate the potential causality links between mucus degradation, its prevention through functional foods, and the control of the inflammatory response of the host.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

Author Contribution

A.L., G.P-V., C.D. and M.M-B. designed the draft of the manuscript. A.L. and G.P-V. wrote the manuscript. J-L.P. made Table 1. A.L. made Tables 2–4. S.L. made Figure 1.

Abbreviations

CAZymes, carbohydrate-active enzymes; Gal, galactose; GalNAc, *N*-acetyl-D-galactosamine; Glc, glucose; GlcNac, *N*-acetyl-D-glucosamine; GH, glycoside hydrolase; GOS, β-galacto-oligosaccharides; HMO, human milk oligosaccharide; IBD, inflammatory bowel diseases; MUC, mucin; Neu5Ac, *N*-acetylneuraminic acid.

References

1   Burger-van Paassen, N., Vincent, A., Puiman, P.J., van der Sluis, M., Bouma, J., Boehm, G. et al. (2009) The regulation of intestinal mucin MUC2 expression by short-chain fatty acids: implications for epithelial protection. *Biochem. J.* **420**, 211–219, https://doi.org/10.1042/BJ20082222

2   Gillois, K., Lévêque, M., Theodorou, V., Robert, H. and Mercier-Bonin, M. (2018) Mucus: an underestimated gut target for environmental pollutants and food additives. *Microorganisms* **6**, 53, https://doi.org/10.3390/microorganisms6020053

3   Talbot, P., Radziwill-Bienkowska, J.M., Kamphuis, J.B.J., Steenkeste, K., Bettini, S., Robert, V. et al. (2018) Food-grade TiO2 is trapped by intestinal mucus *in vitro* but does not impair mucin *O*-glycosylation and short-chain fatty acid synthesis *in vivo*: Implications for gut barrier protection. *J. Nanobiotechnol.* **16**, 53, https://doi.org/10.1186/s12951-018-0379-5

4   da Silva, S., Robbe-Masselot, C., Ait-Belgnaoui, A., Mancuso, A., Mercade-Loubière, M., Salvador-Cartier, C. et al. (2014) Stress disrupts intestinal mucus barrier in rats via mucin *O*-glycosylation shift: prevention by a probiotic treatment. *Am. J. Physiol. Gastrointest. Liver Physiol.* **307**, 420–429, https://doi.org/10.1152/ajpgi.00290.2013

5   Sonnenburg, E.D. and Sonnenburg, J.L. (2014) Starving our microbial self: The deleterious consequences of a diet deficient in microbiota-accessible carbohydrates. *Cell Metab.* **20**, 779–786, https://doi.org/10.1016/j.cmet.2014.07.003

6   Coward, S. and Kaplan, G.G. (2017) IBD in the New World, Old World, and Your World. In *Inflammatory Bowel Disease. Clinical Gastroenterology.* (Cohen, R.D., ed.), pp. 13–27, Humana Press, Cham, https://doi.org/10.1007/978-3-319-53763-4˙2

7   Frank, D.N., Amand, A.L.S., Feldman, R.A., Boedeker, E.C., Harpaz, N. and Pace, N.R. (2007) Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *PNAS* **104**, 13780–13785, https://doi.org/10.1073/pnas.0706625104

8   Png, C.W., Lindén, S.K., Gilshenan, K.S., Zoetendal, E.G., McSweeney, C.S., Sly, L.I. et al. (2010) Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria. *Am. J. Gastroenterol.* **105**, 2420–2428, https://doi.org/10.1038/ajg.2010.281

9   Joossens, M., Huys, G., Cnockaert, M., de Preter, V., Verbeke, K., Rutgeerts, P. et al. (2011) Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. *Gut* **60**, 631–637, https://doi.org/10.1136/gut.2010.223263

10   Torres, J., Mehandru, S., Colombel, J.F. and Peyrin-Biroulet, L. (2017) Crohn's disease. *Lancet* **389**, 1741–1755, https://doi.org/10.1016/S0140-6736(16)31711-1

11   Strugala, V., Dettmar, P.W. and Pearson, J.P. (2008) Thickness and continuity of the adherent colonic mucus barrier in active and quiescent ulcerative colitis and Crohn's disease. *Int. J. Clin. Pract.* **62**, 762–769, https://doi.org/10.1111/j.1742-1241.2007.01665.x

12   Gersemann, M., Becker, S., Kübler, I., Koslowski, M., Wang, G., Herrlinger, K.R. et al. (2009) Differences in goblet cell differentiation between Crohn's disease and ulcerative colitis. *Differentiation* **77**, 84–94, https://doi.org/10.1016/j.diff.2008.09.008

13   Johansson, M.E.V., Gustafsson, J.K., Holmen-Larsson, J., Jabbar, K.S., Xia, L., Xu, H. et al. (2014) Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis. *Gut* **63**, 281–291, https://doi.org/10.1136/gutjnl-2012-303207

14   Johansson, M.E.V. (2014) Mucus layers in inflammatory bowel disease. *Inflamm. Bowel Dis.* **20**, 2124–2131, https://doi.org/10.1097/MIB.0000000000000117

15 El Kaoutari, A., Armougom, F., Gordon, J.I., Raoult, D. and Henrissat, B. (2013) The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat. Rev. Microbiol.* **11**, 497–504, https://doi.org/10.1038/nrmicro3050

16 Lapébie, P., Lombard, V., Drula, E., Terrapon, N. and Henrissat, B. (2019) Bacteroidetes use thousands of enzyme combinations to break down glycans. *Nat. Commun.* **10**, 2043, https://doi.org/10.1038/s41467-019-10068-5

17 Poxton, I.R., Brown, R., Sawyerr, A. and Ferguson, A. (1997) Mucosa-associated bacterial flora of the human colon. *J. Med. Microbiol.* **46**, 85–91, https://doi.org/10.1099/00222615-46-1-85

18 Swidsinski, A., Weber, J., Loening-Baucke, V., Hale, L.P. and Lochs, H. (2005) Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. *J. Clin. Microbiol.* **43**, 3380–3389, https://doi.org/10.1128/JCM.43.7.3380-3389.2005

19 Derrien, M., Vaughan, E.E., Plugge, C.M. and de Vos, W.M. (2004) *Akkermansia municiphila* gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. *Int. J. Syst. Evol. Microbiol.* **54**, 1469–1476, https://doi.org/10.1099/ijs.0.02873-0

20 Hall, A.B., Yassour, M., Sauk, J., Garner, A., Jiang, X., Arthur, T. et al. (2017) A novel *Ruminococcus gnavus* clade enriched in inflammatory bowel disease patients. *Genome Med.* **9**, 103, https://doi.org/10.1186/s13073-017-0490-5

21 Koropatkin, N.M., Cameron, E.A. and Martens, E.C. (2012) How glycan metabolism shapes the human gut microbiota. *Nat. Rev. Microbiol.* **10**, 323–335, https://doi.org/10.1038/nrmicro2746

22 Tailford, L.E., Crost, E.H., Kavanaugh, D. and Juge, N. (2015) Mucin glycan foraging in the human gut microbiome. *Front. Genet.* **6**, 81, https://doi.org/10.3389/fgene.2015.00081

23 Cockburn, D.W. and Koropatkin, N.M. (2016) Polysaccharide degradation by the intestinal microbiota and its influence on human health and disease. *J. Mol. Biol. Acad. Press* **428**, 3230–3252, https://doi.org/10.1016/j.jmb.2016.06.021

24 Ndeh, D. and Gilbert, H.J. (2018) Biochemistry of complex glycan depolymerisation by the human gut microbiota. *FEMS Microbiol. Rev.* **42**, 146–164, https://doi.org/10.1093/femsre/fuy002

25 Brown, H.A. and Koropatkin, N.M. (2021) Host glycan utilization within the Bacteroidetes Sus-like paradigm. *Glycobiology* **31**, 697–706, https://doi.org/10.1093/glycob/cwaa054

26 Laville, E., Perrier, J., Bejar, N., Maresca, M., Esque, J., Tauzin, A.S. et al. (2019) Investigating host microbiota relationships through functional metagenomics. *Front. Microbiol.* **10**, 1286, https://doi.org/10.3389/fmicb.2019.01286

27 Tauzin, A.S., Pereira, M.R., van Vliet, L.D., Colin, P.Y., Laville, E., Esque, J. et al. (2020) Investigating host-microbiome interactions by droplet-based microfluidics. *Microbiome* **8**, 141, https://doi.org/10.1186/s40168-020-00911-z

28 Bode, L. (2009) Human milk oligosaccharides: prebiotics and beyond. *Nutr. Rev.* **67**, S183–191, https://doi.org/10.1111/j.1753-4887.2009.00239.x

29 Fuhrer, A., Sprenger, N., Kurakevich, E., Borsig, L., Chassard, C. and Hennet, T. (2010) Milk sialyllactose influences colitis in mice through selective intestinal bacterial colonization. *J. Exp. Med.* **207**, 2843–2854, https://doi.org/10.1084/jem.20101098

30 He, Y.Y., Liu, S.B., Kling, D.E., Leone, S., Lawlor, N.T., Huang, Y. et al. (2016) The human milk oligosaccharide 2′-fucosyllactose modulates CD14 expression in human enterocytes, thereby attenuating LPS-induced inflammation. *Gut* **65**, 33–46, https://doi.org/10.1136/gutjnl-2014-307544

31 Kurakevich, E., Hennet, T., Hausmann, M., Rogler, G. and Borsig, L. (2013) Milk oligosaccharide sialyl(α2,3)lactose activates intestinal CD11c+ cells through TLR4. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17444–17449, https://doi.org/10.1073/pnas.1306322110

32 Looijer-van Langen, M.A.C. and Dieleman, L.A. (2009) Prebiotics in chronic intestinal inflammation. *Inflamm. Bowel Dis.* **15**, 454–462, https://doi.org/10.1002/ibd.20737

33 Bansil, R. and Turner, B.S. (2018) The biology of mucus: composition, synthesis and organization. *Adv. Drug Deliv. Rev.* **124**, 3–15, https://doi.org/10.1016/j.addr.2017.09.023

34 Reily, C., Stewart, T.J., Renfrow, M.B. and Novak, J. (2019) Glycosylation in health and disease. *Nat. Rev. Nephrol.* **15**, 346–366, https://doi.org/10.1038/s41581-019-0129-4

35 Arike, L. and Hansson, G.C. (2016) The densely *O*-glycosylated MUC2 mucin protects the intestine and provides food for the commensal bacteria. *J. Mol. Biol. Academic Press* **428**, 3221–3229, https://doi.org/10.1016/j.jmb.2016.02.010

36 Larsson, J.M.H., Karlsson, H., Sjövall, H. and Hansson, G.C. (2009) A complex, but uniform *O*-glycosylation of the human MUC2 mucin from colonic biopsies analyzed by nanoLC/MSn. *Glycobiology* **19**, 756–766, https://doi.org/10.1093/glycob/cwp048

37 Ravcheev, D.A., Godzik, A., Osterman, A.L. and Rodionov, D.A. (2013) Polysaccharides utilization in human gut bacterium Bacteroides thetaiotaomicron: Comparative genomics reconstruction of metabolic and regulatory networks. *BMC Genomics* **14**, 873, https://doi.org/10.1186/1471-2164-14-873

38 Glover, J.S., Ticer, T.D. and Engevik, M.A. (2022) Characterizing the mucin-degrading capacity of the human gut microbiota. *Sci. Rep.* **12**, 8456, https://doi.org/10.1038/s41598-022-11819-z

39 Terrapon, N., Lombard, V., Drula, É., Lapébie, P., Al-Masaudi, S., Gilbert, H.J. et al. (2018) PULDB: The expanded database of Polysaccharide Utilization Loci. *Nucleic Acids Res.* **46**, D677–D683, https://doi.org/10.1093/nar/gkx1022

40 Martens, E.C., Chiang, H.C. and Gordon, J.I. (2008) Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host Microbe* **4**, 447–457, https://doi.org/10.1016/j.chom.2008.09.007

41 Crouch, L.I., Liberato, M.v., Urbanowicz, P.A., Baslé, A., Lamb, C.A., Stewart, C.J. et al. (2020) Prominent members of the human gut microbiota express endo-acting *O*-glycanases to initiate mucin breakdown. *Nat. Commun.* **11**, 4017, https://doi.org/10.1038/s41467-020-17847-5

42 Barbeyron, T., Brillet-Guéguen, L., Carré, W., Carrière, C., Caron, C., Czjzek, M. et al. (2016) Matching the diversity of sulfated biomolecules: Creation of a classification database for sulfatases reflecting their substrate specificity. *PloS ONE* **11**, e0164846, https://doi.org/10.1371/journal.pone.0164846

43 Luis, A.S., Jin, C., Pereira, G.V., Glowacki, R.W.P., Gugel, S.R., Singh, S. et al. (2021) A single sulfatase is required to access colonic mucin by a gut bacterium. *Nature* **598**, 332–337, https://doi.org/10.1038/s41586-021-03967-5

44   Luis, A.S., Baslé, A., Byrne, D.P., Wright, G.S.A., London, J.A., Jin, C. et al. (2022) Sulfated glycan recognition by carbohydrate sulfatases of the human gut microbiota. *Nat. Chem. Biol.* **18**, 841–849, https://doi.org/10.1038/s41589-022-01039-x

45   Pudlo, N.A., Urs, K., Kumar, S.S., German, J.B., Mills, D.A. and Martens, E.C. (2015) Symbiotic human gut bacteria with variable metabolic priorities for host mucosal glycans. *mBio* **6**, e01282–15, https://doi.org/10.1128/mBio.01282-15

46   Donaldson, G.P., Chou, W.C., Manson, A.L., Rogov, P., Abeel, T., Bochicchio, J. et al. (2020) Spatially distinct physiology of *Bacteroides fragilis* within the proximal colon of gnotobiotic mice. *Nat. Microbiol.* **5**, 746–756, https://doi.org/10.1038/s41564-020-0683-3

47   Ottman, N., Davids, M., Suarez-Diez, M., Boeren, S., Schaap, P.J., dos Santos, V.A.P.M. et al. (2017) Genome-scale model and omics analysis of metabolic capacities of *Akkermansia muciniphila* reveal a preferential mucin-degrading lifestyle. *Appl. Environ. Microbiol.* **83**, e01014–17, https://doi.org/10.1128/AEM.01014-17

48   Shin, J., Noh, J.R., Chang, D.H., Kim, Y.H., Kim, M.H., Lee, E.S. et al. (2019) Elucidation of *Akkermansia muciniphila* probiotic traits driven by mucin depletion. *Front. Microbiol.* **10**, 1137, https://doi.org/10.3389/fmicb.2019.01137

49   Kostopoulos, I., Aalvink, S., Kovatcheva-Datchary, P., Nijsse, B., Bäckhed, F., Knol, J. et al. (2021) A Continuous Battle for Host-Derived Glycans Between a Mucus Specialist and a Glycan Generalist *in vitro* and *in vivo*. *Front. Microbiol.* **12**, 632454, https://doi.org/10.3389/fmicb.2021.632454

50   Liu, X., Zhao, F., Liu, H., Xie, Y., Zhao, D. and Li, C. (2021) Transcriptomics and metabolomics reveal the adaption of *Akkermansia muciniphila* to high mucin by regulating energy homeostasis. *Sci. Rep.* **11**, 9073, https://doi.org/10.1038/s41598-021-88397-z

51   Crost, E.H., Tailford, L.E., le Gall, G., Fons, M., Henrissat, B. and Juge, N. (2013) Utilisation of Mucin Glycans by the Human Gut Symbiont *Ruminococcus gnavus* Is Strain-Dependent. *PloS ONE* **8**, e76341, https://doi.org/10.1371/journal.pone.0076341

52   Turroni, F., Milani, C., van Sinderen, D. and Ventura, M. (2011) Genetic strategies for mucin metabolism in Bifidobacterium bifidum PRL2010: An example of possible human-microbe co-evolution. *Gut Microbes* **2**, 183–189, https://doi.org/10.4161/gmic.2.3.16105

53   Kosciow, K. and Deppenmeier, U. (2020) Characterization of three novel β-galactosidases from *Akkermansia muciniphila* involved in mucin degradation. *Int. J. Biol. Macromol.* **149**, 331–340, https://doi.org/10.1016/j.ijbiomac.2020.01.246

54   Pichler, M.J., Yamada, C., Shuoker, B., Alvarez-Silva, C., Gotoh, A., Leth, M.L. et al. (2020) Butyrate producing colonic Clostridiales metabolise human milk oligosaccharides and cross feed on mucin via conserved pathways. *Nat. Commun.* **11**, 3285, https://doi.org/10.1038/s41467-020-17075-x

55   Wu, H., Rebello, O., Crost, E.H., Owen, C.D., Walpole, S., Bennati-Granier, C. et al. (2021) Fucosidases from the human gut symbiont *Ruminococcus gnavus*. *Cell. Mol. Life Sci.* **78**, 675–693, https://doi.org/10.1007/s00018-020-03514-x

56   Fan, S., Zhang, H., Chen, X., Lu, L., Xu, L. and Xiao, M. (2016) Cloning, characterization, and production of three α-l-fucosidases from *Clostridium perfringens* ATCC 13124. *J. Basic Microbiol.* **56**, 347–357, https://doi.org/10.1002/jobm.201500582

57   Shuoker, B., Pichler, M.J., Jin, C., Hiroka, S., Wu, H., Gascueña, A.M. et al. (2022) Sialidases and Fucosidases of *Akkermansia muciniphila* are crucial for growth on mucin and nutrient sharing with mucus-associated gut bacteria. *bioRxiv*, https://doi.org/10.1101/2022.09.10.507281

58   Rahfeld, P., Wardman, J.F., Mehr, K., Huff, D., Morgan-Lang, C., Chen, H.M. et al. (2019) Prospecting for microbial α-*N*-acetylgalactosaminidases yields a new class of GH31 *O*-glycanase. *J. Biol. Chem.* **294**, 16400–16415, https://doi.org/10.1074/jbc.RA119.010628

59   Tailford, L.E., Owen, C.D., Walshaw, J., Crost, E.H., Hardy-Goddard, J., le Gall, G. et al. (2015) Discovery of intramolecular *trans*-sialidases in human gut microbiota suggests novel mechanisms of mucosal adaptation. *Nat. Commun.* **6**, 7624, https://doi.org/10.1038/ncomms8624

60   Robinson, L.S., Lewis, W.G. and Lewis, A.L. (2017) The sialate *O*-acetylesterase EstA from gut Bacteroidetes species enables sialidase-mediated cross-species foraging of 9-*O*-acetylated sialoglycans. *J. Biol. Chem.* **292**, 11861–11872, https://doi.org/10.1074/jbc.M116.769232

61   Guo, B.S., Zheng, F., Crouch, L., Cai, Z.P., Wang, M., Bolam, D.N. et al. (2018) Cloning, purification and biochemical characterisation of a GH35 beta-1,3/beta-1,6-galactosidase from the mucin-degrading gut bacterium *Akkermansia muciniphila*. *Glycoconj. J.* **35**, 255–263, https://doi.org/10.1007/s10719-018-9824-9

62   Kosciow, K. and Deppenmeier, U. (2019) Characterization of a phospholipid-regulated β-galactosidase from *Akkermansia muciniphila* involved in mucin degradation. *Microbiol. Open* **8**, e00796, https://doi.org/10.1002/mbo3.796

63   Henderson, I.R., Czeczulin, J., Eslava, C., Noriega, F. and Nataro, J.P. (1999) Characterization of pic, a secreted protease of *Shigella flexneri* and enteroaggregative *Escherichia coli*. *Infect. Immun.* **67**, 5587–5596, https://doi.org/10.1128/IAI.67.11.5587-5596.1999

64   Nakjang, S., Ndeh, D.A., Wipat, A., Bolam, D.N. and Hirt, R.P. (2012) A novel extracellular metallopeptidase domain shared by animal Host-Associated mutualistic and pathogenic microbes. *PloS ONE* **7**, e30287, https://doi.org/10.1371/journal.pone.0030287

65   Noach, I., Ficko-Blean, E., Pluvinage, B., Stuart, C., Jenkins, M.L., Brochu, D. et al. (2017) Recognition of protein-linked glycans as a determinant of peptidase activity. *Proc. Natl. Acad. Sci. U.S A.* **114**, E679–E688, https://doi.org/10.1073/pnas.1615141114

66   Shon, D.J., Malaker, S.A., Pedram, K., Yang, E., Krishnan, V., Dorigo, O. et al. (2020) An enzymatic toolkit for selective proteolysis, detection, and visualization of mucin-domain glycoproteins. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 21299–21307, https://doi.org/10.1073/pnas.2012196117

67   Nesta, B., Valeri, M., Spagnuolo, A., Rosini, R., Mora, M., Donato, P. et al. (2014) SslE Elicits Functional Antibodies That Impair In Vitro Mucinase Activity and In Vivo Colonization by Both Intestinal and Extraintestinal *Escherichia coli* Strains. *PLoS Pathog.* **10**, e1004124, https://doi.org/10.1371/journal.ppat.1004124

68   Luo, Q., Kumar, P., Vickers, T.J., Sheikh, A., Lewis, W.G., Rasko, D.A. et al. (2014) Enterotoxigenic *Escherichia coli* secretes a highly conserved mucin-degrading metalloprotease to effectively engage intestinal epithelial cells. *Infect. Immun.* **82**, 509–521, https://doi.org/10.1128/IAI.01106-13

69   Shon, D.J., Fernandez, D., Riley, N.M., Ferracane, M.J. and Bertozzi, C.R. (2022) Structure-guided mutagenesis of a mucin-selective metalloprotease from *Akkermansia muciniphila* alters substrate preferences. *J. Biol. Chem.* **298**, 101917, https://doi.org/10.1016/j.jbc.2022.101917

70   Taleb, V., Liao, Q., Narimatsu, Y., García-García, A., Compañón, I., Borges, R.J. et al. (2022) Structural and mechanistic insights into the cleavage of clustered *O*-glycan patches-containing glycoproteins by mucinases of the human gut. *Nat. Commun.* **13**, 4324, https://doi.org/10.1038/s41467-022-32021-9

71   Bardoel, B.W., Hartsink, D., Vughs, M.M., de Haas, C.J.C., van Strijp, J.A.G. and van Kessel, K.P.M. (2012) Identification of an immunomodulating metalloprotease of *Pseudomonas aeruginosa* (IMPa). *Cell. Microbiol.* **14**, 902–913, https://doi.org/10.1111/j.1462-5822.2012.01765.x

72   Pluvinage, B., Ficko-Blean, E., Noach, I., Stuart, C., Thompson, N., McClure, H. et al. (2021) Architecturally complex *O*-glycopeptidases are customized for mucin recognition and hydrolysis. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2019220118, https://doi.org/10.1073/pnas.2019220118

73   Hews, C.L., Tran, S.L., Wegmann, U., Brett, B., Walsham, A.D.S., Kavanaugh, D. et al. (2017) The StcE metalloprotease of enterohaemorrhagic *Escherichia coli* reduces the inner mucus layer and promotes adherence to human colonic epithelium *ex vivo*. *Cell. Microbiol.* **19**, e12717, https://doi.org/10.1111/cmi.12717

74   Trastoy, B., Naegeli, A., Anso, I., Sjögren, J. and Guerin, M.E. (2020) Structural basis of mammalian mucin processing by the human gut *O*-glycopeptidase OgpA from *Akkermansia muciniphila*. *Nat. Commun.* **11**, 4844, https://doi.org/10.1038/s41467-020-18696-y

75   Almeida, A., Nayfach, S., Boland, M., Strozzi, F., Beracochea, M., Shi, Z.J. et al. (2021) A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* **39**, 105–114, https://doi.org/10.1038/s41587-020-0603-3

76   Tasse, L., Bercovici, J., Pizzut-Serin, S., Robe, P., Tap, J., Klopp, C. et al. (2010) Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. *Genome Res.* **20**, 1605–1612, https://doi.org/10.1101/gr.108332.110

77   Ladevéze, S., Tarquis, L., Cecchini, D.A., Bercovici, J., André, I., Topham, C.M. et al. (2013) Role of glycoside phosphorylases in mannose foraging by human gut bacteria. *J. Biol. Chem.* **288**, 32370–32383, https://doi.org/10.1074/jbc.M113.483628

78   Almeida, A., Mitchell, A.L., Boland, M., Forster, S.C., Gloor, G.B., Tarkowska, A. et al. (2019) A new genomic blueprint of the human gut microbiota. *Nature* **568**, 499–504, https://doi.org/10.1038/s41586-019-0965-1

79   Pérez, J.C. (2021) Fungi of the human gut microbiota: roles and significance. *Int. J. Med. Microbiol.* **311**, 151490, https://doi.org/10.1016/j.ijmm.2021.151490

80   van Bueren, A.L., Mulder, M., van Leeuwen, S. and Dijkhuizen, L. (2017) Prebiotic galacto-oligosaccharides activate mucin and pectic galactan utilization pathways in the human gut symbiont *Bacteroides thetaiotaomicron*. *Sci. Rep.* **7**, 40478, https://doi.org/10.1038/srep40478

81   Chia, L.W., Hornung, B.V.H., Aalvink, S., Schaap, P.J., de Vos, W.M., Knol, J. et al. (2018) Deciphering the trophic interaction between *Akkermansia muciniphila* and the butyrogenic gut commensal *Anaerostipes caccae* using a metatranscriptomic approach. *Antonie van Leeuwenhoek, Int. J. General Mol. Microbiol.* **111**, 859–873, https://doi.org/10.1007/s10482-018-1040-x

82   Li, L., Zhang, X., Ning, Z., Mayne, J., Moore, J.I., Butcher, J. et al. (2018) Evaluating *in vitro* culture medium of gut microbiome with orthogonal experimental design and a metaproteomics approach. *J. Proteome Res.* **17**, 154–163, https://doi.org/10.1021/acs.jproteome.7b00461

83   Li, A., Laville, E., Tarquis, L., Lombard, V., Ropartz, D., Terrapon, N. et al. (2020) Analysis of the diversity of the glycoside hydrolase family 130 in mammal gut microbiomes reveals a novel mannoside-phosphorylase function. *Microb. Genom.* **6**, mgen000404, https://doi.org/10.1099/mgen.0.000404

84   Erickson, A.R., Cantarel, B.L., Lamendella, R., Darzi, Y., Mongodin, E.F., Pan, C. et al. (2012) Integrated metagenomics/metaproteomics reveals human host-microbiota signatures of Crohn's disease. *PloS ONE* **7**, e49138, https://doi.org/10.1371/journal.pone.0049138

85   Nordman, H., Davies, J.R., Herrmann, A., Karlsson, N.G., Hansson, G.C. and Carlstedt, I. (1997) Mucus glycoproteins from pig gastric mucosa: identification of different mucin populations from the surface epithelium. *Biochem. J.* **326**, 903–910, https://doi.org/10.1042/bj3260903

86   Padra, M., Adamczyk, B., Benktander, J., Flahou, B., Skoog, E.C., Padra, J.T. et al. (2018) *Helicobacter suis* binding to carbohydrates on human and porcine gastric mucins and glycolipids occurs via two modes. *Virulence* **9**, 898–918, https://doi.org/10.1080/21505594.2018.1460979

87   Robbe, C., Capon, C., Maes, E., Rousset, M., Zweibaum, A., Zanetta, J.P. et al. (2003) Evidence of regio-specific glycosylation in human intestinal mucins: Presence of an acidic gradient along the intestinal tract. *J. Biol. Chem.* **278**, 46337–46348, https://doi.org/10.1074/jbc.M302529200

88   Jin, C., Kenny, D.T., Skoog, E.C., Padra, M., Adamczyk, B., Vitizeva, V. et al. (2017) Structural diversity of human gastric mucin glycans. *Mol. Cell. Proteomics* **16**, 743–758, https://doi.org/10.1074/mcp.M117.067983

89   Robbe, C., Capon, C., Coddeville, B. and Michalski, J.C. (2004) Structural diversity and specific distribution of *O*-glycans in normal human mucins along the intestinal tract. *Biochem. J.* **384**, 307–316, https://doi.org/10.1042/BJ20040605

90   Katayama, T., Fujita, K. and Yamamoto, K. (2005) Novel bifidobacterial glycosidases acting on sugar chains of mucin glycoproteins. *J. Biosci. Bioeng.* **99**, 457–465, https://doi.org/10.1263/jbb.99.457

91   Gagnon, M., Zihler Berner, A., Chervet, N., Chassard, C. and Lacroix, C. (2013) Comparison of the Caco-2, HT-29 and the mucus-secreting HT29-MTX intestinal cell models to investigate Salmonella adhesion and invasion. *J. Microbiol. Methods* **94**, 274–279, https://doi.org/10.1016/j.mimet.2013.06.027

92   Nason, R., Büll, C., Konstantinidi, A., Sun, L., Ye, Z., Halim, A. et al. (2021) Display of the human mucinome with defined *O*-glycans by gene engineered cells. *Nat. Commun.* **12**, 4070, https://doi.org/10.1038/s41467-021-24366-4

93   Wang, Y., Kim, R., Sims, C.E. and Allbritton, N.L. (2019) Building a thick mucus hydrogel layer to improve the physiological relevance of in vitro primary colonic epithelial models. *CMGH* **8**, 653.e5–655.e5, https://doi.org/10.1016/j.jcmgh.2019.07.009

94   Zhao, A., Qin, H., Sun, M., Tang, M., Mei, J., Ma, K. et al. (2021) Chemical conversion of human epidermal stem cells into intestinal goblet cells for modeling mucus-microbe interaction and therapy. *Sci. Adv.* **7**, eabb2213, https://doi.org/10.1126/sciadv.abb2213

95   Gielen, F., Hours, R., Emond, S., Fischlechner, M., Schell, U. and Hollfelder, F. (2016) Ultra-high-throughput-directed enzyme evolution by absorbance-activated droplet sorting (AADS). *Proc. Natl. Acad. Sci. U. S. A.* **113**, E7383–E7389, https://doi.org/10.1073/pnas.1606927113

96   Dagkesamanskaya, A., Langer, K., Tauzin, A.S., Rouzeau, C., Lestrade, D., Potocki-Veronese, G. et al. (2018) Use of photoswitchable fluorescent proteins for droplet-based microfluidic screening. *J. Microbiol. Methods* **147**, 59–65, https://doi.org/10.1016/j.mimet.2018.03.001

97   Villa, M.M., Bloom, R.J., Silverman, J.D., Durand, H.K., Jiang, S., Wu, A. et al. (2020) Interindividual variation in dietary carbohydrate metabolism by gut bacteria revealed with droplet microfluidic culture. *mSystems* **5**, e00864–19, https://doi.org/10.1128/mSystems.00864-19

98   Kamphuis, J.B.J., Mercier-Bonin, M., Eutamène, H. and Theodorou, V. (2017) Mucus organisation is shaped by colonic content; A new view. *Sci. Rep.* **7**, 8527, https://doi.org/10.1038/s41598-017-08938-3

99   Desai, M.S., Seekatz, A.M., Koropatkin, N.M., Kamada, N., Hickey, C.A., Wolter, M. et al. (2016) A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. *Cell* **167**, 1339.e21–1353.e21, https://doi.org/10.1016/j.cell.2016.10.043

100 Marcobal, A., Barboza, M., Sonnenburg, E.D., Pudlo, N., Martens, E.C., Desai, P. et al. (2011) Bacteroides in the infant gut consume milk oligosaccharides via mucus-utilization pathways. *Cell Host Microbe* **10**, 507–514, https://doi.org/10.1016/j.chom.2011.10.007

101 Holma, R., Juvonen, P., Asmawi, M.Z., Vapaatalo, H. and Korpela, R. (2002) Galacto-oligosaccharides stimulate the growth of bifidobacteria but fail to attenuate inflammation in experimental colitis in rats. *Scand. J. Gastroenterol.* **37**, 1042–1047, https://doi.org/10.1080/003655202320378239

102 Belzer, C. (2022) Nutritional strategies for mucosal health: the interplay between microbes and mucin glycans. *Trends Microbiol.* **30**, 13–21, https://doi.org/10.1016/j.tim.2021.06.003

103 Elison, E., Vigsnaes, L.K., Rindom Krogsgaard, L., Rasmussen, J., Sorensen, N., McConnell, B. et al. (2016) Oral supplementation of healthy adults with 2′-*O*-fucosyllactose and lacto-*N*-neotetraose is well tolerated and shifts the intestinal microbiota. *Br. J. Nutr.* **116**, 1356–1368, https://doi.org/10.1017/S0007114516003354

104 Angeloni, S., Ridet, J.L., Kusy, N., Gao, H., Crevoisier, F., Guinchard, S. et al. (2005) Glycoprofiling with micro-arrays of glycoconjugates and lectins. *Glycobiology* **15**, 31–41, https://doi.org/10.1093/glycob/cwh143