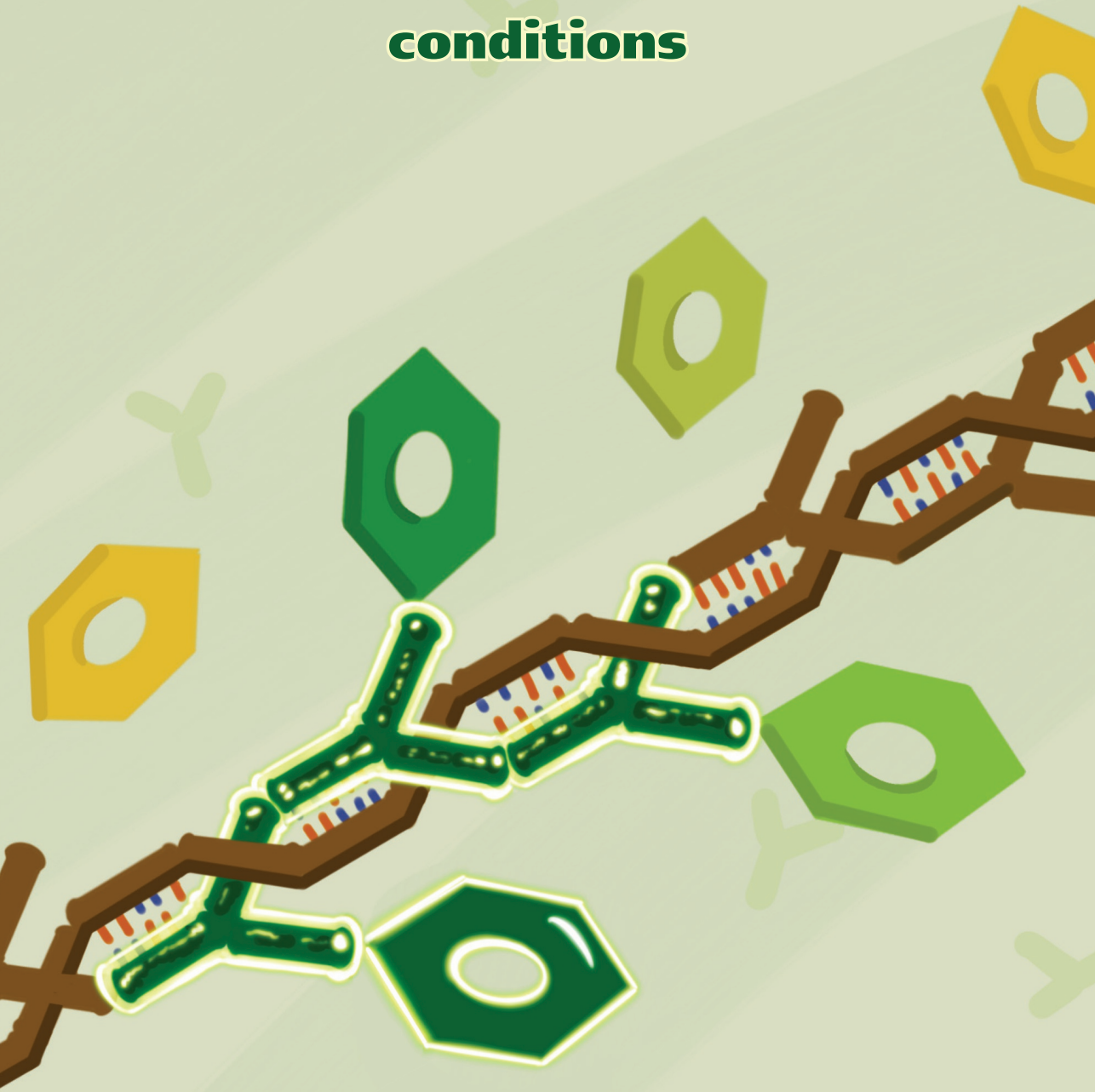
Unravelling serpin regulation in *Bifidobacterium longum* NCC 2705; from mechanisms to production conditions



Stéphane Duboux

Propositions

1. The trans-acting sBLO-161 sRNA plays a global post-transcriptional regulatory role in *Bifidobacterium longum* subsp. *longum*.
   (this thesis)

2. A tailored synbiotic that delivers a serpin-producing *Bifidobacterium longum* NCC 2705 to the colon effectively reduces excessive colonic inflammation.
   (this thesis)

3. The 'art' of wine making should capitalize much more on the increasing scientific understanding of the functional diversity of fungal [1] and bacterial [2] starter cultures
   [1] Querol A et al. *Adv Food Nutr Res. 2018;85:177-210*
   [2] Vicente J et al. *Int J Food Microbiol. 2022 Aug 16;375:109726*

4. In the global warming context, management of soil microbiome diversity is a necessity to achieve temperature increase reduction goals [3]
   [3] Shankar Bhattacharyya et al. *Science of the Total Environment. 2022 Apr 1;815:152928*

5. The skills acquired through the practice of music conducting are an asset for pursuing scientific research

6. Nuanced scientific argumentation is not effective in convincing idealistic or dogmatic policy makers

Propositions belonging to the thesis, entitled

Unravelling serpin regulation in *B. longum* NCC 2705; from mechanisms to production conditions

Stéphane Duboux
Wageningen, 09 December 2022

Unravelling serpin regulation in *Bifidobacterium longum* NCC 2705; from mechanisms to production conditions

Stéphane Duboux

Thesis committee

Promotor

Prof. Dr M. Kleerebezem

Personal chair, Host-Microbe Interactomics Group
Wageningen University & Research

Co-promotors

Dr. Annick Mercenier, NutriLeads, Wageningen, The Netherlands
Dr. Biljana Bogicevic, Nestlé Research, Lausanne, Switzerland

Other members

Prof. T. Abee, Wageningen University & Research
Prof. D. van Sinderen, University College Cork, Ireland
Prof. J. Kok, Groningen University, The Netherlands
Dr. T.D. Leser, Chr. Hansen A/S, Copenhagen, Denmark

This research was conducted under the auspices of the Graduate School VLAG (Advanced studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences)

Unravelling serpin regulation in *Bifidobacterium longum* NCC 2705; from mechanisms to production conditions

Stéphane Duboux

Thesis

submitted in fulfilment of the requirements for the degree of doctor

at Wageningen University

By the authority of the Rector Magnificus,

Prof. Dr. A.P.J. Mol,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Friday 9th of December 2022

at 4:00 p.m. in the Omnia Auditorium

Stéphane Duboux

Unravelling serpin regulation in *Bifidobacterium longum* NCC 2705; from mechanisms to production conditions

260 pages

PhD thesis, Wageningen University, Wageningen, The Netherlands (2022)

With references, with summary in English

ISBN: 978-94-6447-362-9

DOI: https://doi.org/10.18174/575256

Table of contents

CHAPTER 1

General introduction

S. Duboux [1,2]

[1] Nestlé Research, Route du Jorat 57, CH 1000 Lausanne 26, Switzerland
[2] Host-Microbe Interactomics Group, Wageningen University & Research, De Elst 1, 6708WD Wageningen, The Netherlands

Bifidobacteria, an important genus in early infancy

The origin of the *Bifidobacterium* taxon and current diversity. The first description of bifidobacteria was presented by the pediatrician H. Tissier in 1899. He discovered an abundant anaerobic bacterium with a bifurcated (bifid) shape in the feces of breast fed babies, which he initially named *Bacillus bifidus* [1]. Later he even proposed to use this bacteria in the treatment of intestinal infections [2]. After more than a century of research, bifidobacteria have been isolated from a wide range of ecological niches, such as oral cavity, sewage, human blood, fermented or raw milk, as well as the hindgut of birds and insects (Figure 1). However, the vast majority of currently described bifidobacterial species originate from the gastrointestinal tract of human and other mammals [3]. Today, the *Bifidobacterium* genus encompass 98 published species [4], proposed to be organized in seven phylogenetic groups: the *B. adolescentis* group, the *B. asteroides* group, the *B. boum* group, the *B. longum* group, the *B. bifidum* group, the *B. pseudolongum* group, and the *B. pullorum* group [5] (Figure 1).

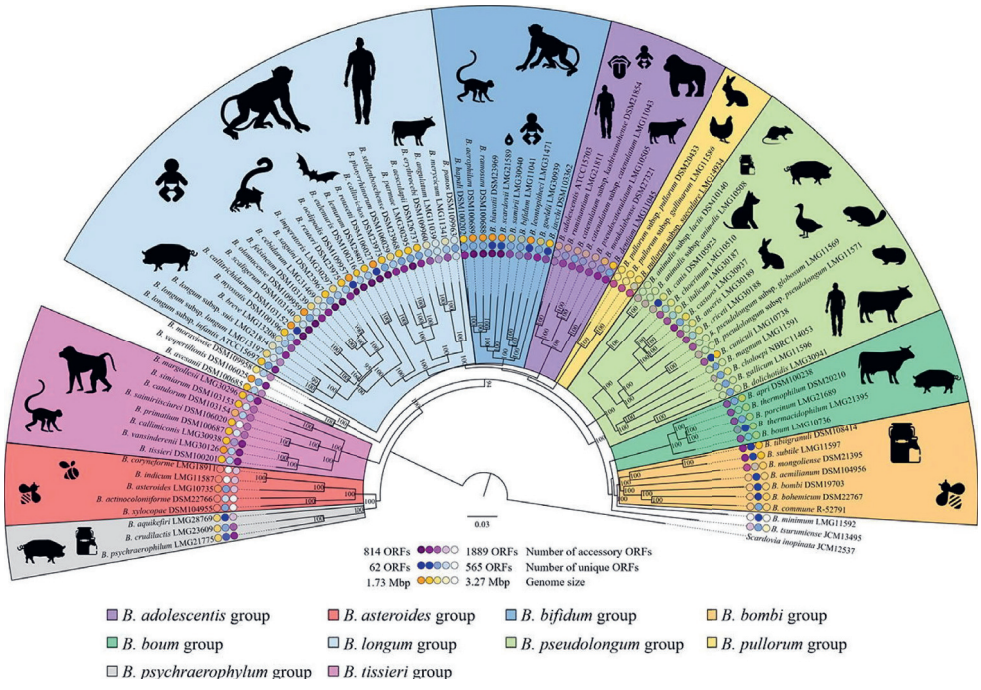


Figure 1. Phylogenetic tree of the genus *Bifidobacterium* based on the concatenation of the 169 amino acid sequences representing the *Bifidobacterium* core genome. The phylogenetic tree was generated by the maximum-likelihood method, and COG sequences of *Scardovia inopinata* JCM 12537 shared with bifidobacterial species were used as an outgroup. Bootstrap percentages above 50 are shown at node points based on 1000 replicates of the phylogenetic tree. The outer circle represents the number of accessory and unique genes as well as the genome size of each bifidobacterial type strain, while the ecological origins of bifidobacteria per each phylogenetic group are reported beside the phylogenetic tree. This figure was published in "The genus *bifidobacterium*: From genomics to functionality of an important component of the mammalian gut microbiota, Comput Struct Biotechnol J 19, 1472-1487, Alessandri et al. 2021, ©Elsevier" [3].

Introduction

Lifespan distribution of species found in human and related carbohydrate metabolism. Strains belonging to the *B. adolescentis* group (*B. catenulatum*, *B. pseudocatenulatum*, *B. dentium*, *B. adolescentis*), the *B. longum* group (*B. longum*, *B. breve*), the *B. pullorum* group (*B. bifidum)* and the *B. pseudolongum* group (*B. animalis*), are commonly detected in the feces of healthy human individuals. Amongst those, the last two species have been proposed not to belong to the autochthonous microbiota of the human gastrointestinal tract (GIT) because they were not detected in mucosa-associated samples. Hidalgo-Cantabrana et al. proposed that the presence of *B. animalis* in faecal samples is possibly explained by the consumption of probiotic dairy products or food supplements, whereas *B. dentium* is a typical inhabitant of the human oral cavity and may thereby also be encountered in faecal samples [6].

Bifidobacterial species such as *B. longum, B. bifidum* and *B. breve* are particularly abundant in the neonate human GIT, being predominant in the breast-fed newborn [7]. The early life predominancy of those taxons, especially *B. longum* subsp. *infantis*, was proposed to be linked to their capacity to hydrolyze and metabolize several of the indigestible human milk oligosaccharides (HMOs) [8,9]. Notably, strain level differences have been reported for the capacity of *B. longum subsp. infantis* to metabolize specific HMOs, where certain strains preferentially consumed sialylated HMOs, while others preferably fermented fucosylated HMOs [10]. Furthermore, probably as a consequence of evolutionary diversification, different *Bifidobacterium* species encode variable, and relatively complex HMOs import and fermentation strategies that include syntrophic relations among the members of this genus. For example, *B. bifidum* possesses two extracellular exo-α-sialidases that can release sialic acid from sialylated HMOs outside of the cell [11], which can subsequently be imported and metabolized by *B. breve* [12,13], creating a commensalistic relation between these *Bifidobacterium* species.

Bifidobacteria relative abundance decreases after the breast-feeding period, as upon introduction of solid food other bacterial taxa appear and take over part of the gastrointestinal niche [14]. However, it is proposed that in early adulthood, the bifidobacterial community stabilizes and remains relatively stable until old age, suggesting a relevant role for this genus during the entire lifespan [3,15]. Their long-term residence in the human gastrointestinal tract is believed to be at least in part explained

by their capacity to effectively breakdown and metabolize diverse diet-derived carbohydrate substrates [16]. For example, bifidobacterial strains belonging to *B. longum* subsp. *longum* or *B. adolescentis* harbor the enzymatic machinery necessary to degrade various food-derived carbohydrates that are not digested by mammalian hosts. *B. longum* subsp. *longum* strains has been recognized to effectively metabolize carbohydrates found in human milk such as lactose [17] or Lacto-N-Tetraose (LNT) [18], but also appear to be particularly well equipped to degrade, import and consume food derived polysaccharides, which is prominently reflected by the large arsenal of glucosidases encoded by the genome of *B. longum* subsp. *longum* NCC 2705 [19]. Conversely, *B. adolescentis* does not harbor genes related to the consumption of complex human milk oligosaccharides but encodes a large repertoire of glycan hydrolases that can target food related substrates, such as resistant starch [20]. The role of these bifidobacteria in the digestion of food derived carbohydrates is further supported by the observation that bifidobacterial communities are stimulated by the consumption of cereal fibers, such as arabinogalactans [21] or arabinoxylans derived oligosaccharides [22], or inulin containing vegetables, such as chicory, artichoke, onion, leek or salsify [23,24].

The beneficial roles of bifidobacteria in early infancy. The first human gut colonizers, which bifidobacteria belong to, are proposed to occupy the intestinal niche and to produce a set of molecules that inhibits or promotes the presence of other bacterial populations at a later stage [25]. Initially, bifidobacteria were proposed to inhibit pathogens through the stimulation of the immune system or the production of antibacterial peptides and quorum-sensing inhibitors [6]. In addition, the short chain fatty acids (SCFAs), mainly acetic acid and lactic acid, produced by bifidobacteria play an important role in early life, as they contribute to lowering of the intestinal pH, which prevents the outgrowth of enteropathogens [26] and provides substrates for butyrate production by other bacteria [27]. Importantly, intestinal microbially produced SCFAs, and in particular butyric acid, serve to fuel epithelial cells, eliciting beneficial effects on metabolism [28], inflammation and thereby play a protective role in a variety of diseases [29,30]. Several studies have highlighted the mutualistic relationship between bifidobacteria and other gut commensals. For example, in a gnotobiotic mouse model, *Bacteroides thetaiotaomicron* was able to extend its polysaccharide utilization capacity in presence of NCC 2705. Authors proposed that this enhanced transcription of genes

implicated in mannoside and xyloside metabolism resulted from a change in gastro-intestinal carbohydrate distribution induced by the presence of *B. longum* NCC 2705 [31]. Another study demonstrated a cross-feeding mechanism between *B. longum* and *Eubacterium rectale*, were *B. longum* produced acetate from arabinoxylans, allowing *E. rectale* to produce butyrate, overall explaining the bifidogenic and butyrogenic effect of diet-derived arabinoxylans [32]. As well, co-cultivation of NCC 2705 with *Bacteroides caccae* enhanced arabinogalactan degradation, through the combined action of their extracellular β-galactanases and β-galactosidases. This likely catalyzed the formation of smaller oligosaccharides and overall promoted the growth of the *B. longum* strain [33]. Finally, several secondary metabolites produced by bifidobacteria are believed to promote microbial diversity in the gut. For example, bifidobacterial exopolysaccharides may serve as a growth substrate for other gut commensals [34,35]. Additionally, B-vitamins, which can be produced by certain *Bifidobacterium* strains [36], are proposed to promote diversity and stability of the gut microbiome, as a significant fraction of microbial gut communities (>20%) are composed of auxotrophic species whose viability is strictly dependent on acquiring one or more B-vitamins from diet and/or prototrophic microbes [37].

The presence of bifidobacteria in the gut during early human life was proposed to play a critical role in the maturation of the newborn's immune system, and several studies have shown associations between low abundance of members of this genus in the infant gut to later in life immune disorders such as atopy [38] and asthma [39]. Moreover, early-life *Bifidobacterium* dominance is associated with reduced risks of obesity later in life [40]. High bifidobacteria levels in infancy is as well associated with a reduced risk of acquiring antimicrobial resistant bacteria, based on the purported *Bifidobacterium* capacity to suppress antimicrobial resistance-carrying taxa [41]. These associations have inspired the development and clinical evaluation of bifidobacterial probiotic strains targeting a relatively large range of benefits, including the prevention of diarrhea (antibiotic or *Clostridium difficile* associated), infections (e.g., *Helicobacter pylori*), Inflammatory Bowel Disease (IBD), Irritable Bowel Syndrome (IBS), Necrotizing Enterocolitis (NEC) and allergies. Although there is an abundant list of positive clinical results associated with the use of *Bifidobacterium* probiotics for the treatment and prevention of several of these diseases, only few strains have been demonstrated to provide a consistent health effect in a congruent set of clinical trials

[6], as for example *B. animalis* subsp. *lactis* BB-12 for the prevention of acute gastroenteritis in children [42]. The efforts to develop bifidobacterial probiotics continues, and in May 2022 there were 201 clinical trials registered in ClinicalTrials.gov that include interventions with bifidobacterial strains in young children (up to 17 months), targeting a wide array of health conditions. We can therefore anticipate that the range of bifidobacterial probiotic strains that are accompanied with rigorously demonstrated beneficial effects on health may expand in the future.

The model organism *B. longum* NCC 2705

Origin and general metabolic features. *B. longum* NCC 2705 represents a prominent *Bifidobacterium* model strain as it is one of the most studied. NCC 2705 was isolated from infant feces at Nestlé Research (Lausanne, Switzerland) and was subsequently deposited in 1996 in the Nestlé Culture Collection (NCC; Lausanne, Switzerland). It was the first fully sequenced bifidobacterial strain and its genome was published in 2002 [16].

The initial genome analysis revealed that the strain has the capacity to produce 19 amino acids from NH_4 and other major biosynthetic precursors proposing that this strain has very few amino acid auxotrophies. To date, experimental verification of the functionality of the identified amino acid synthesis pathways is lacking, warranting dedicated investigations by for example single amino acid omission experiments. In addition, NCC 2705 is predicted to synthetize several important co-factors like folic acid, thiamin and nicotinate. Genome analysis also indicated that the strain is likely auxotroph for other co-factors like riboflavin, biotin, cobalamin, pantothenate and pyridoxine [16]. Finally, the genome of *B. longum* NCC 2705 appeared to be particularly adapted to utilize a diverse range of carbohydrates, as more than 8.5% of its overall predicted gene repertoire was associated with carbohydrate transport and metabolism. Moreover, the genome encodes more than 40 predicted glycosyl hydrolases, including 2 xylanases, 9 arabinosidases, 2 α-galactosidases, neopullanase, isomaltase, maltase, inulinase (β-fructofuranosidase), 4 β-galactosidases, 3 β-glucosidases, 3 hexosaminidases, and 3 α-mannosidases [16], which are foreseen to support the degradation of a variety of complex carbohydrates.

The carbohydrate metabolism of *B. longum* NCC 2705. Following initial genome-based description, the carbohydrate metabolism of *B. longum* NCC 2705 has been the subject of numerous studies establishing the exact function of specific genes and highlighting some of their tight regulation of expression. An important study by Parche et al. showed that NCC 2705 was able to grow on arabinose, fructose, galactose, glucose, ribose, xylose, maltose, melibiose, sucrose, lactose, raffinose, arabinogalactan and raftilose [19]. A detailed prediction of the sugar transport system encoded by NCC 2705 was pursued, identifying a total of 19 import systems, including no less than 13 ATP-binding cassette (ABC) transporters. Transcription of 13 of these

transport function encoding genes was shown to be modulated by growth on lactose, raffinose, maltose or fructo-oligosaccharides (FOS). Remarkably, 9 of the 13 transporter systems appeared to be induced during growth on FOS [19], and a specific ABC transporter system (*fruEKFG*; *BL0033-0036*) was demonstrated to mediate fructose import in this strain. The substrate binding protein of this system (FruE) was shown to not only have affinity for fructose, but also could bind ribose and xylose, and its expression was strongly suppressed by the presence of glucose [43,44]. In a subsequent study, it was demonstrated that a subset of these carbohydrates (glucose, fructose, mannose, xylose, ribose and galactose) were all catabolized via the so-called "bifid-shunt" [44]. Bifidobacteria lack the fructose 6-phosphokinase enzyme (EC 2.7.1.11) that directly converts fructose 6-phosphate into fructose 1,6-biphoshate and is used by lactic acid bacteria to channel glucose into the conventional glycolysis pathway [45]. The "bifid-shunt" is instead organized around the pivotal xylulose-5-phosphate/fructose-6-phosphate phosphoketolase enzyme (F6PPK) (encoded by *BL0959* in NCC 2705) [46] that can similarly convert fructose-6-phosphate to erythrose-4-phosphate and acetyl phosphate or xylulose-5-phosphate into glyceraldehyde-3-phosphate and acetyl phosphate. This unique setup allows bifidobacteria to produce more ATP from 1 mole of glucose than the fermentative pathways of lactic acid acid bacteria (2.5 vs 1-2 moles of ATP, respectively) [46].

Overall, the carbohydrate metabolism of NCC 2705 reflects its adaptation to the human gut environment, with a particular adaptation to the infant gut [16]. Prominent in that adaptation is the strain's demonstrated preference for lactose over glucose, which is illustrated by the lactose-mediated repression of the glucose transporter GlcP (encoded by the *glcP* gene; *BL1631*) that has the highest affinity for glucose but can also import mannose and galactose [17]. Although it was initially annotated as a FOS transport and utilization system [19], the *BL1638-1641* operon was further demonstrated to mediate the import and metabolism of the human milk oligosaccharide lacto-N-tetraose (LNT), including a gene encoding a specialized lacto-N-biose phosphorylase (*LNBP; BL1641*) [18]. The NCC 2705 genome also harbors a gene cluster proposed to be involved in the import and metabolization of host N-glycans derived mannosylated oligosaccharides. This cluster includes a specialized β-mannosidase (encoded by *BL1333*) that removes the mannosyl moieties from these oligosaccharides [47]. NCC 2705 does not only metabolize milk or host-derived glycans, but also has the capacity

to use various food-derived carbohydrates. For example, the strain encodes a sucrose phosphorylase (*scrP; BL0536*) and a β-fructofuranosidase (*cscA; BL0105*) that both displayed sucrase activity. Transcription of the *scrP* gene was induced when the cells were grown on sucrose and raffinose, whereas it was repressed in presence of glucose. Besides sucrose, the *cscA* encoded enzyme also hydrolyses raftilose and inulin. Based on these findings, Kullin et al. proposed ScrP to cleave intracellular sucrose and ScsA to specifically cleave shorter fructose oligosaccharides (sFOS) [48]. The *BL0257* encoded extracellular glycan hydrolase displaying endo-β-1,4-galactanase (GH53) activity [33] was initially proposed to be part of a lactose or FOS transport operon [19], but was subsequently demonstrated to specifically hydrolyse β1->4 and β1->3 galactans, like those contained in potatoes [49] or cereals [33]. Finally, the strain produces an extracellular endo-1,4-β-xylanase (*xynD; BL1543*) that together with its neighboring gene encoding an extracellular arabinofuranosidase (*BL1544*), was strongly induced during growth on arabinoxylan oligosaccharides (AXOS) [50].

Overall, the carbohydrate metabolism of NCC 2705 is amongst the best studied in bifidobacteria. Since carbohydrate catabolism is the most predominant source of energy (ATP) in bifidobacteria, the members of this genus need to strictly regulate carbon import and metabolic fluxes [51]. Although the studies mentioned above describe some of the details of this rigorous and complex regulation in NCC 2705 (Table 1), to date we do not have a complete understanding of the underlying gene regulation mechanisms for this strain.

Table 1: Carbohydrate related functions previously shown to be differentially regulated in *B. longum* NCC 2705.

Genes ID	Predicted Function	Regulation	Other Evidences	Reference
BL0033-36	ABC transport system for fructose and fos	BL0033-36: ↑ fos; ↑ fructose, ribose, xylose BL0033: ↓ glucose	BL0033 binds to fructose with the highest affinity	Parche et al. 2007 Lui et al 2011 Wei et al. 2012
BL0141-0146	Import (ABC system) & metabolization of maltose & arabinose	BL0141-0146: ↑ maltose BL0141: ↓ FOS BL0146: ↑ FOS		Parche et al. 2007
BL0257-0264	Import (ABC system) & metabolization of galactans & galacto-oligosaccharides	BL0262: ↑ FOS	BL0257 demonstrated to be a GH53 endo-beta-1,4-galactanase, with a signal peptide (extracellular)	Wang et al. 2020 Hinz et al. 2005 Parche et al. 2007
BL0420-0425	Import (ABC system) & metabolization of xylans or arabinans	BL0425: ↑ FOS		Parche et al. 2007
BL1163-1171	Import (ABC system) & metabolization of arabinose containing oligosaccharides	BL1164, 65, 70: ↑ lactose BL1164, 65: ↑ FOS		Parche et al. 2007
BL1327-1336	Import (ABC system) & metabolization of mannosylated glycan (host N-glycan)	BL1327: ↑ lactose BL1327: ↓ maltose & FOS	BL1333 is a specialized B-mannosidase involved in N-glycan degradation	Cordeiro et al. 2019 Parche et al. 2007
BL1518-1526	Import (ABC system) & metabolization of raffinose	BL1521-23: ↑ raffinose		Parche et al. 2007
BL1638-BL1641	Import (ABC system) & metabolization of Lacto-N-tetraose (LNT)	BL1638-40: ↑ FOS	BL1641 is a lacto-N-biose phosphoryalase (LNBP)	Kitoka et al. 2005 Parche et al. 2007
BL1691-96	ABC transport system for arabinose	BL1694, 96: ↑ FOS BL1696: ↑ raffinose		Parche et al. 2007
BL0105-0107	Import (Major Facilitator System) & metabolization of sucrose and FOS	BL0105-0107: ↑ FOS	BL0105 is a beta-fructofuranosidase demonstrated to hydrolyze raftilose, inulin and sucrose	Kullin et al. 2006 Parche et al. 2007
BL0534-0536	Import (Major Facilitator System) & metabolization of sucrose and raffinose	BL0536: ↑ sucrose, raffinose BL0536: ↓ glucose	BL0536 is a sucrose phosphorylase with a demonstrated sucrase activity	Kullin et al. 2006
BL0976-0978	Import (Major Facilitator System) & metabolization of galacto-oligosaccharides (including lactose)	BL0978: ↑ lactose, maltose, FOS		Parche et al. 2007
BL1631	Major Facilitator transport System for glucose, mannose and galactose	BL1631: ↓ lactose, raffinose, maltose, FOS	BL1631 is proton symporter with binding preference for glucose>mannose>galactose	Parche et al. 2006 Parche et al. 2007
BL0410	Major Intrinsic Protein family for glycerol-like	BL0410: ↓ FOS		Parche et al. 2007

The niche factors and effector molecules supporting NCC 2705 persistence and health-beneficial effects in the gut. Various studies have revealed some of the mechanisms employed by *B. longum* NCC 2705 to support its establishment and survival in the gastro-intestinal tract (Figure 2 A). Yuan et al. investigated the strain's proteome adaptation during *in vivo* growth (incubation in a dialysis tubing implanted within a rabbit intestine), identifying a higher presence of the bile salt hydrolase (*Bsh, BL0796*) and elongation factor Tuf (*tuf; BL1097*) [52]. The Bsh enzyme is well known to catalyze the deconjugation of primary bile acids commonly conjugated with taurine or glycine residues, releasing the amino acid conjugate and the secondary bile acids. The resulting secondary bile acids are then exported by different transporters such as the cholate transporter Ctr (*ctr*; BL1102) [53] or the bile inducible Major Facilitator Superfamily (MFS) transporter (*BL0920*) [54]. Interestingly, the Tuf protein and the enolase protein (*eno; BL1022*) were concomitantly increased during co-incubation of NCC 2705 with gut epithelial cells ($CaCo_2$) *in vitro*. When recombinantly produced, both proteins inhibited NCC 2705 adhesion to epithelial cells [55] suggesting that these proteins are exposed on the NCC 2705 cell surface and can thereby facilitate its binding to epithelial cells. Until recently, it was not clear how these typical intracellular proteins could be exposed to the cell surface. However, Nishiyama et al. recently demonstrated that *B. longum* NCC 2705 produces extracellular vesicles that contain several cytoplasmic proteins, such as the GroEL (*BL002*), Tuf, transaldolase (Tal encoded by *BL0715*), and heat shock protein 20 (Hsp20 encoded by *BL0576*). The same study also established that recombinantly produced GroEL and Tal displayed high-affinity binding to mucins *in vitro*. This led the authors to propose that extracellular vesicles could serve as export vehicles for adhesion molecules in NCC 2705 [56]. Besides the surface exposure of so-called "moon-lighting" cytoplasmic enzymes, NCC 2705 encodes polymorphic surface exposed fimbriae (*BL0675*) that display high affinity binding to porcine mucins and thereby could contribute to the strain adhesion to the gastrointestinal mucosa [57] (Figure 2 A).

B. longum NCC 2705 is also equipped with different systems to cope with stress conditions encountered in the gastrointestinal tract. The Bsh-mediated deconjugation of bile acids and their export to the environment (see above), reduces the surfactant stress posed by the primary bile acids [53,54]. Next to this, NCC 2705 was demonstrated to intrinsically display a relevant level of resistance to oxidative stress [58], which was

suggested to be mediated by an alkyl hydroperoxide reductase (*ahpC; BL0615*) [59,60]. Finally, NCC 2705 encodes a surface exposed serine protease inhibitor (*BL0108*) [16], which was shown to form covalent products with pancreatic and neutrophil elastases and was proposed to protect the strain against the damaging effects of host-derived proteases [61] (Figure 2 A).

The proposed probiotic properties and the role of the underlying effector molecules of *B. longum* NCC 2705 were investigated in several preclinical studies (Figure 2 B). The antipathogenic activity of this strain was demonstrated in both *in vitro* and *in vivo* models. For example, acetic acid production by NCC 2705 inhibited *E. coli* O157 *in vitro*, which translated into a protective effect against this enteropathogen in a murine model. A mutant of NCC 2705 lacking the *BL0033* encoded fructose binding protein FruE, not only reduced the strain capacity to catabolize fructose and produce acetate when grown on this sugar, but also led to a reduced protection against *E. coli* O157 infection *in vivo* [26]. In another study, Mundi et al. demonstrated that a NCC 2705 cell-free culture supernatant suppressed the expression of two *Campylobacter jejuni* genes that are implicated in the virulence of this species (invasion antigen, CiaB and the flagellin A, FlaA). The authors also demonstrated the involvement of a *Campylobacter jejuni* quorum-sensing mechanism since the effects of the culture supernatant was abolished in a *LuxS* mutant [62].

The immunomodulatory role of *B. longum* NCC 2705 was demonstrated in several studies. Anti-inflammatory effects of NCC 2705 were established initially *in vitro* by demonstrating that the strain could inhibit LPS-induced NF-κB-mediated inflammation in HT-29 reporter cells, reducing the production of LPS-induced pro-inflammatory cytokines (IL-8, TNF-α, Cox-2, and ICAM-1) [63]. Furthermore, in a mouse model of gluten sensitivity, NCC 2705 was demonstrated to dampen gliadin-induced immunopathology. This protective effect was shown to be serpin-mediated, since a serpin knock-out strain failed to protect gliadin sensitized mice [64]. Moreover, this study confirmed that pure recombinantly produced NCC 2705 serpin inhibits human neutrophil elastase [61,64], which is released at the sites of intestinal inflammation [65], supporting a role of the *B. longum* NCC 2705 serpin in dampening innate immune responses. Finally, *B. longum* NCC 2705 was shown to induce a T-helper 1 (TH1) type of immune response in mono-associated gnotobiotic mice and was proposed to play a role in restoring an appropriate TH1/TH2 immune balance that may contribute to the

prevention of the onset of allergic diseases [66]. In a follow up study, the same authors proposed that the TH1 immunostimulatory property was linked to the presence of unmethylated CpG motifs in the DNA of *B. longum* NCC 2705, since these DNA sequences induced the production of TH1 related cytokines (such as MCP-1 and TNF-α) in RAW 264.7 macrophages [67] (Figure 2 B).

The NCC 2705 was recently tested in a safety clinical trial performed in celiac disease and non-celiac gluten sensitive patients. Results revealed that the strain is safe for human consumption and well tolerated. In the same study, appearance of NCC 2705 in duodenal aspirates was associated to a concomitant increase in serpin concentration, providing a mechanistic basis for the potential use of the strain in this population [68]. Although the pre-clinical effects of NCC 2705 have not yet been translated to similar and clinically relevant effects in human studies, the fact remains that *B. longum* NCC 2705 is one of the best characterized bifidobacterial strain. A wealth of studies has shed light on its carbohydrate metabolism and the molecules driving its persistence and capacity to interact with the host (Figure 2) leading to a relatively broad understanding of its functioning. This knowledge remains to be complemented by studies focusing on the mechanisms underlying the regulation of expression of these diverse functions.

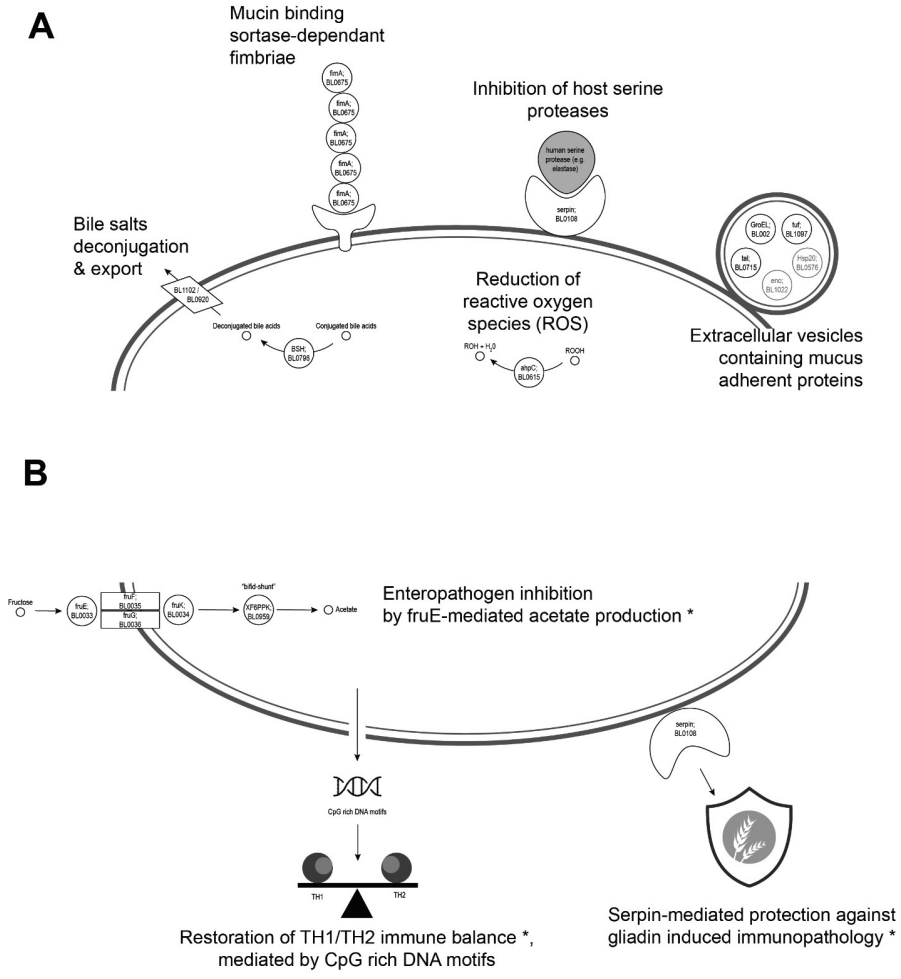


Figure 2: Experimentally established niche factors (panel A) and effector molecules (panel B) of *B. longum* NCC 2705. Mechanisms demonstrated in *in vivo* studies are highlighted by a star (*).

Production processes and their potential implications for the bioactivity of probiotics.

Probiotic delivery formats. The most widely accepted definition of probiotics was proposed by the FAO/WHO in 2002 [69] and was slightly modified by an expert panel constituted by the International Scientific Association for Probiotics and Prebiotics (ISAPP) in 2014 [70]: "Probiotics are live microorganisms that, when administered in adequate amounts, confer a health benefit on the host". This definition implies that probiotics should be alive at the time of consumption, whatever the format they are delivered in. Today, probiotics are delivered in two main formats, in refrigerated (mostly dairy) fermented products and shelf-stable dried powders. Many differences in the production process of those two probiotic product formats exist, including the medium composition used for growth (dairy based vs synthetic medium) and the downstream processing to prepare the final product (presence vs absence of supernatant, refrigeration vs drying, etc.). Studies that directly compare those two production formats and the impact they have on the physiology of the probiotic strain, including how these could affect its intestinal survival and/or efficacy to elicit the desired health effects are lacking.

The stresses imposed by manufacturing. Production of concentrated stable probiotic powders involves different steps: a) medium preparation and sterilization; b) fermentation; c) centrifugation and supernatant removal; d) addition of drying protective agents; e) drying [71]. Throughout these steps, a bifidobacterial strain can be exposed to a relatively large range of stresses, including oxidative stress (exposure to oxygen during growth or drying), nutrient starvation (depending on the ingredients nature and levels used), pH and osmotic stress (depending on the pH regulation strategy), mechanical stress (during centrifugation and other manipulations), as well as cold (during freeze-drying) or heat (during spray-drying) stresses [72]. Bifidobacterial strains possess a range of mechanisms enabling them to survive these stresses, and production strategies aiming to activate these mechanisms have already been suggested in 2008 [73] and subsequently tested [74]. As an example, several studies have proposed that exposing bifidobacteria to a pH stress during cultivation is beneficial as this elicits stress response reactions that increase resistance to low pH conditions and various other environmental stresses through cross protection mechanisms [75,76]. Low pH adaptation in bifidobacteria may at least partially be mediated by cell wall

modifications, which could affect cell surface composition and protein exposure (e.g., adhesin) [75,77].

The potential impact of nutrient formulations during fermentation. Bifidobacteria have a variety of strain-specific nutritional requirements in terms of amino acids, cofactors, vitamins and carbohydrates. Recently, genome-based metabolic models have been constructed to ease the identification of strain-specific growth-limiting factors [78], which may facilitate the development of strain-specific synthetic media in the future. However, at present, fulfilling those nutritional requirements often relies on the addition of yeast autolysates (yeast extracts) and hydrolyzed proteins (hereafter peptones) to provide a generic source of nitrogen, vitamins and other micronutrients [79]. Yeast extracts and peptones are complex ingredients, and their source or manufacturing process influence the level of amino acids, vitamins, minerals and residual carbohydrates they contain, and thus may affect the composition of the final medium [80-82]. In certain cases, cane and beet molasses can be used as the carbon source for growth, and their geographical origin and levels of purity may impact the performance and yield of the fermentation [71]. Optimization of the type, quantity and quality of these growth medium ingredients is therefore needed to ensure high yields in bacterial biomass, which are required to establish an economically viable probiotic production process [71]. To further increase biomass yield during fermentation, pH regulation can be applied during industrial probiotic production to prevent premature growth stagnation induced by acidification [83].

All the above suggest that variations in manufacturing processes are likely to impact the metabolism and physiology of probiotic *Bifidobacterium* strains. To develop more reliable manufacturing processes that do not only guarantee high cell counts, but also ensure the consistency of the targeted health-benefit, there is a need to better understand the implications of probiotic manufacturing procedures that go beyond the standard yield and stability readouts.

Outline of the thesis

Discrepancies in clinical trial outcomes have been described for probiotic strains belonging to the *Lactobacillaceae* family. Different clues obtained by research on strains belonging to this family suggest that variations in manufacturing conditions could in part be responsible for those discrepant outcomes, since they may modify the presence and activity of important probiotic effector molecules exposed at the surface of the bacteria [84]. Similarly, inconsistencies in bifidobacterial clinical trial outcomes have been observed, but the potential contribution of manufacturing conditions on these variable outcomes has not been studied at the level of their functional molecular make-up to date. In **Chapter 2**, we summarize the current knowledge in this area, focusing on the potential effect that manufacturing might have on the expression or presence of bifidobacterial niche factors and effector molecules.

The other chapters of the thesis focus on *B. longum* NCC 2705 and serpin, using this tandem as an exemplary model to better understand how production processes may influence both the production of this important bifidobacterial effector molecule and the metabolism and overall physiology of the strain. In **chapter 3**, we investigate the carbohydrate-mediated regulation of serpin production in *B. longum* NCC 2705, highlighting that it differs from the previously described *B. breve* serpin control mechanism. In this chapter, we also propose the existence in this strain of a regulation circuit that resembles common glucose catabolite repression. **Chapter 4** deciphers the role of different pH regulation strategies during biomass production in a scaled-down fermentation process. We demonstrate that depending on the selected pH regulation strategy, glucose catabolite repression might be relieved resulting not only in an increased serpin production but also contributing to an enhanced *B. longum* NCC 2705 resistance to low pH. Using a fluorescent reporter system that we implemented for use in bifidobacteria, we measure in **chapter 5** the activity of specific promoters that control carbohydrate utilization operons. By this approach, we decipher the activation patterns of these promoters when the NCC 2705 strain is grown on different single or mixed sugars. Furthermore, by coupling the fluorescent protein promoter-probe approach to flow cytometry, we establish that exposure to different mixes of carbohydrates elicits different levels of glucose-mediated catabolite repression in *B. longum* NCC 2705.

The results of chapters 3, 4, and 5 demonstrate that in *B. longum* NCC 2705, the production of serpin is tightly linked to the carbohydrates it metabolizes. The data we obtained suggest that the physiological state of the strain is significantly modified by the sugar substrate used. This is further explored in **chapter 6**, where we perform a combination of transcriptome analyses and a range of physiological *in vitro* assays to identify the changes (beyond serpin production) that are induced by growth on glucose or galactose as sole carbon source. This study reveals two highly distinct phenotypes when the strain is grown on these sugars and demonstrates that depending on the carbon source, the bioactivity of the strain is significantly modulated. In this chapter we also revisit previously described global regulatory proteins that are likely to be at least in part involved in the orchestration of these modulations.

The last experimental chapter (**chapter 7**) aims to further complete our understanding of the regulation mechanisms underlying the observed carbohydrate-based adaptations. We show in this chapter that serpin production in *B. longum* NCC 2705 is not subject to transcriptional regulation, which inspired us to investigate the possible role of regulatory sRNAs in post-transcriptional regulation. We identify a potential *trans*-acting sRNA that is predicted to bind and inhibit the translation of not only serpin mRNA but also a range of additional functions in *B. longum* NCC 2705 that are critical in carbon source adaptation. We hereby complement our understanding of the catabolite repression-like regulatory mechanism in this strain that encompasses not only the involvement of regulatory proteins but also a multitarget regulatory sRNA.

Chapter 8 summarizes the main findings of this thesis and discusses potential future research directions. We first revise our initially proposed serpin regulation model in *B. longum* NCC 2705. Then, we especially focus on bringing together all available evidence that demonstrates the existence of a catabolite repression mechanism in bifidobacteria and propose a global regulatory model involving both regulator-protein and sRNA. In this chapter, we also suggest ways forward for the development of *B. longum* NCC 2705 in the management of celiac disease and IBD, two inflammatory diseases which could benefit from the beneficial role of the strain and the serpin it produces

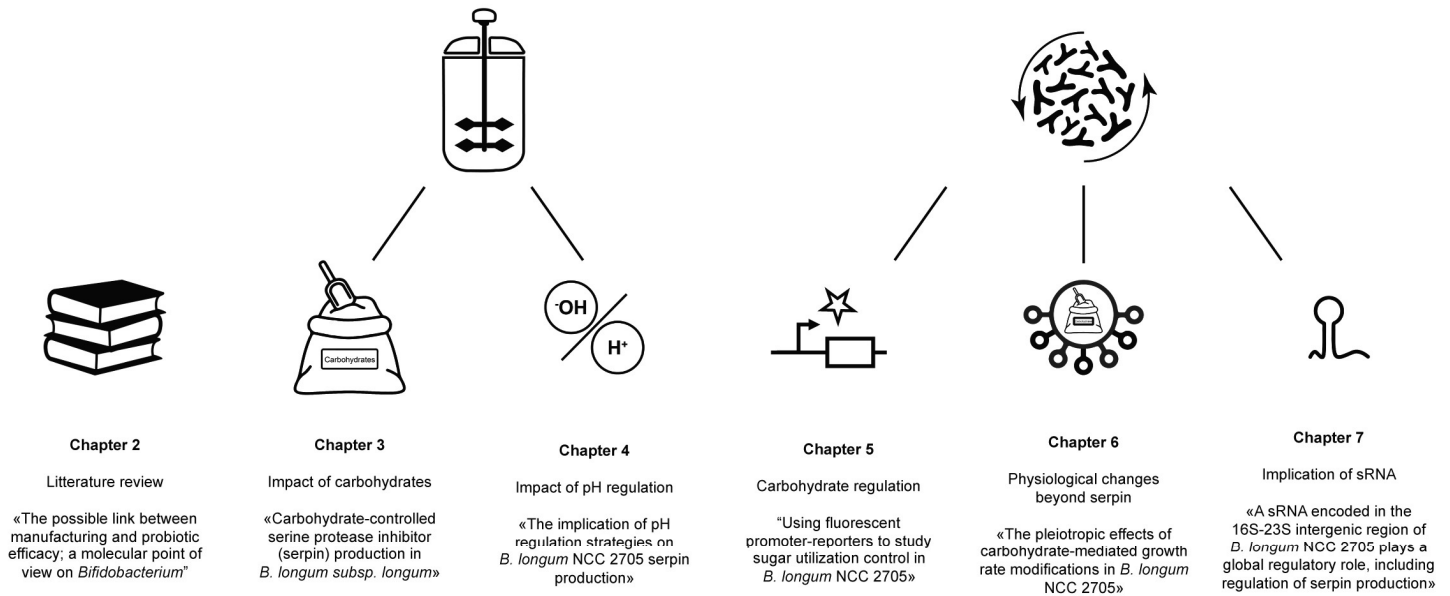


Figure 3: Outline of the thesis summarized in a graphical abstract.

CHAPTER 2

2

The possible link between manufacturing and probiotic efficacy; a molecular point of view on *Bifidobacterium*

S. Duboux [1,2] | M. Van Wijchen [1,2] | M. Kleerebezem [2]

[1] Nestlé Research, Route du Jorat 57, CH 1000 Lausanne 26, Switzerland
[2] Host-Microbe Interactomics Group, Wageningen University & Research, De Elst 1, 6708WD Wageningen, The Netherlands

This chapter was published as Duboux, S., Van Wijchen, M. & Kleerebezem, M. The Possible Link Between Manufacturing and Probiotic Efficacy; a Molecular Point of View on Bifidobacterium. Front Microbiol 12, 812536, doi:10.3389/fmicb.2021.812536 (2021)

Abstract

Probiotics for food or supplement use have been studied in numerous clinical trials, addressing a broad variety of diseases and conditions. However, discrepancies were observed in the clinical outcomes stemming from the use of *Lactobacillaceae* and bifidobacteria strains. These differences are often attributed to variations in the clinical trial protocol like trial design, included target population, probiotic dosage, or outcome parameters measured. However, a contribution of the methods used to produce the live bioactive ingredients should not be neglected as a possible additional factor in the observed clinical outcome variations. It is well established that manufacturing conditions play a role in determining the survival and viability of probiotics, but much less is known about their influence on the probiotic molecular composition and functionality. In this review, we briefly summarize the evidence obtained for *Lacticaseibacillus rhamnosus* GG and *Lactiplantibacillus plantarum* WCFS1, highlighting that expression and presence of probiotic niche factor (NF) and/or effector molecules (EM) may be altered during production of those two well-characterized *Lactobacillaceae* probiotic strains. Subsequently, we summarize in more depth what is the present state of knowledge about bifidobacterial probiotic NF and EM; how their expression may be modified by manufacturing related environmental factors and how that may affect their biological activity in the host. This review highlights the importance of gathering knowledge on probiotic NF and EM, to validate them as surrogate markers of probiotic functionality. We further propose that monitoring of validated NF and/or EM during production and/or in the final preparation could complement viable count assessments that are currently applied in industry. Overall, we suggest that implementation of molecular level quality controls (i.e., based on validated NF and EM), could provide mode of action based *in vitro* tests contributing to better control the health-promoting reliability of probiotic products.

Introduction

Initially formulated by the World Health Organization in 2002 [85] and slightly corrected by experts in the field in 2014 [70], probiotics are today defined as "live microorganisms that, when administered in adequate amounts, confer a health benefit to the host". Overall, probiotic bacteria for food or supplement use (mainly *Lactobacillaceae* & bifidobacteria) have been studied in a large number of clinical trials, targeting a wide array of diseases and conditions [86].

Two distinguishable classes of health benefits are attributed to probiotics: a "general" class of effects that groups beneficial effects exerted by various well-studied microbial species; and a "strain-specific" class of effects that are expected to be driven by specific probiotics strains. An expert panel convened in 2013 has acknowledged those two classes, concluding that "general" benefits such as "creating a more favorable gut environment" and "supporting a healthy digestive tract" (regrouping a diversity of clinical end points such as diarrhoea, antibiotic-associated diarrhoea (AAD), gut transit, abdominal pain, bloating and necrotizing enterocolitis) are displayed by a large number of probiotic strains representing various commonly studied species. The mechanisms of action supporting those "general" probiotic beneficial effects (e.g. probiotic and/or microbiome mediated SCFA production, regulation of intestinal transit, competitive exclusion of pathogens) are similarly believed to be shared by a large number, if not all, of the probiotic strains [70]. Furthermore, "general" benefits (e.g., AAD prevention) provided by the commonly used *Lacticaseibacillus rhamnosus* GG strain have be shown to be relatively consistent throughout different clinical trials in children [42].

In contrast, "strain-specific" benefits are defined as effects that are likely exerted by a limited number of strains, such as "prevention of allergic disease", "downregulation of inflammation", "enhancement of anti-infection activities" or "support of specific organs health" (e.g. reproductive tract, lungs) [70]. Those beneficial effects are believed to be driven by specific molecules present within or at the surface of the probiotic bacterial cells [87-89]. In the last decade, it was shown that a range of molecules produced by probiotic contribute to their robustness and stress tolerance, supporting their survival and establishment when they transit through the gastro-intestinal tract (i.e., so-called niche factors; NF). In addition, various probiotic effector molecules (EM) have been

identified to drive *in situ* host-microbe interactions, thus determining the specific health benefit of different probiotic strains [90]. Disentangling the NF or EM role of specific probiotic molecules is not trivial, especially when adhesive-like phenotypes are affected. Adhesion to the host cells or to the intestinal mucus can be regarded as a factor promoting the bacterial colonization but could as well contribute to the exposure of different structures present on the bacterial cell envelope.

Variability in health effects is not uncommonly observed in clinical trials with probiotic for food or supplement use [91-93]. The inconsistency in results is usually attributed to variations in the design of the clinical studies, including differences in dosage of the probiotic, selection of different target population (i.e., inclusion and exclusion factors at enrolment), powering of the studies according to the primary and secondary objectives, duration of the studies, probiotic delivery format and schedule, data collection and further analysis performed. Indeed, these factors have been suggested to explain part of the discrepancies observed in the reported clinical health-outcomes [94]. Furthermore, probiotics need to exert their effects in the complex microbiome. Inter-individual microbiota variability represent hence a challenge in ensuring probiotic effect consistency in different populations, and new stratification as well as personalized nutrition approaches have been recently proposed to improve the situation [95]. Moreover, the way the probiotic strains themselves are produced and formulated is often not well-described in clinical trial studies, while it could play an important role in the health-promoting efficacy that a product elicits. This is well illustrated by the discrepancies observed in randomized clinical trials using *L. rhamnosus* GG targeting the prevention of allergic disease, which are summarized by Segers & Lebeer [96]. Initially, Kalliomäki and colleagues showed in a landmark study that *L. rhamnosus* GG treatment significantly lowered the risk of eczema in young children belonging to families with a history of atopic disease [97,98]. However, in a subsequent attempt to reproduce this Finnish study protocol, Kopp and colleagues failed to detect similar beneficial effects in a German cohort [99]. Population (Finnish vs Germans) and dosage differences (2^{E10} vs 1^{E10} CFU daily) are potential confounding factor in those two studies, but it is important to note that the source (and possibly the manufacturing process) of *L. rhamnosus* GG in the studies by Kalliomäki et al. [97,98] and Kopp et al. [99] was different, which deserves attention because it may have as well contributed to the differences in clinical outcomes [100]. At present, probiotic

manufacturing procedures remain largely unexplored as a potential source of variation in probiotic clinical trials outcomes and hence deserves to be studied in more details [101,102], especially in the light of the increasing knowledge about specific NF and EM that play a role in the efficacy of intestinal delivery and health promotion following consumption of the product.

Production of dried probiotic supplements consists generally of a) a series of fermentations of different scale where bacterial biomass is produced using specific media and growth conditions, b) a centrifugation step to remove the culture supernatant and concentrate the biomass, c) a mixing step where protectants are added, followed by d) a drying step [71]. Throughout these manufacturing stages, probiotics encounter a range of different stress conditions, including variations in temperature, acid exposure, osmotic and oxidative stress, all of which can modulate their physiology and molecular composition. These modulations may impact their survival during manufacturing as well as their fitness during gastrointestinal tract transit [73,103]. We suggest that not only the expression of NF (e.g., proteins that contribute to robustness and stress tolerance) can be affected by the production conditions, but also effector molecule expression levels may differ, leading to variable presence of molecules that have been shown to mediate the health-benefit elicited by the strain. The required presence of NF and/or EM is further supported by the fact that probiotic cells rendered metabolically inactive by the mean of heat-treatment can still elicit beneficial health effects [104], highlighting the potential limitation of using live cells enumeration alone to ascertain efficacy of probiotic preparations. Therefore, monitoring the expression of validated NF as well as EM during probiotic manufacturing may enable better control of product properties at a molecular level, which goes beyond the traditionally used colony forming units (CFU), and could contribute to an increased robustness of clinical outcomes (Figure 1).

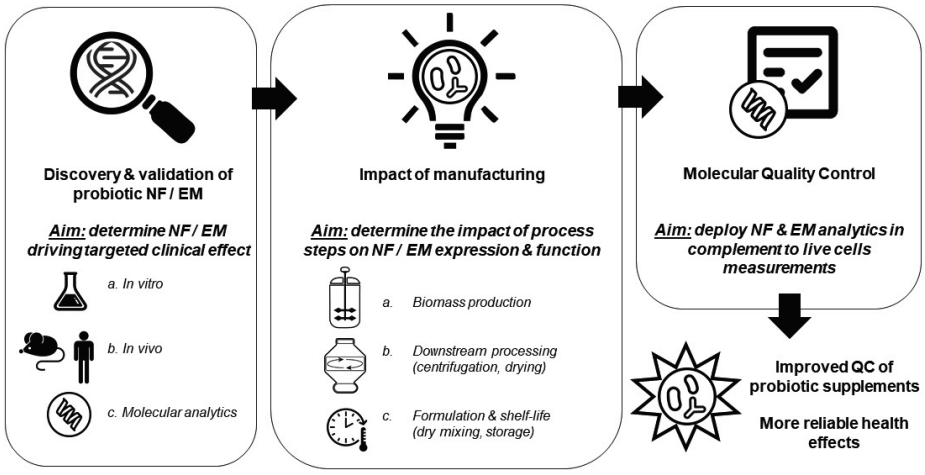


Figure 1 : Stepwise approach leveraging probiotic niche factors (NF) and effector molecules (EM) knowledge to improve probiotic supplements consistency. As a first step, a robust link between the presence of probiotic NF or EM and the desired clinical outcome needs to be established. This step will also allow the development of sets of molecular analytics that can be used in the subsequent steps. Then, the different manufacturing steps need to be evaluated to understand their contribution to NF or EM expression and function. Finally, the NF / EM molecular analytics developed earlier can be used as quality control (QC), complementing traditional CFU / live cells measurements, which overall should enable production of probiotic supplements with increased consistency in their attributed health benefits

In this review, we first briefly highlight that expression and presence of probiotic EM may be altered during production using two well-characterized examples among the *Lactobacillaceae* probiotic, i.e., *Lacticaseibacillus rhamnosus* GG and *Lactiplantibacillus plantarum* WCFS1 (Figure 2). Subsequently, we focus with more depth on bifidobacteria and their probiotic NF and EM, summarizing what is the present state of knowledge about bifidobacterial NF and EM; how their expression may be modified by manufacturing related environmental factors and how that may affect their biological activity in the host (Table 1). Overall, this review highlights the importance of gathering knowledge on probiotic NF and EM, to validate them as surrogate markers of probiotic functionality. We further hypothesize that understanding the dynamics of these molecules during production could contribute to better control the health-promoting reliability of dried probiotic products. Hence, NF / EM based *in vitro* assays represent molecular-level quality controls that adds to the limited information coming from viable count assessments that are currently applied in industry and research.

Modification of *L. rhamnosus* GG effector molecules by processing

L. rhamnosus GG is among the probiotic strains with the best described set of EM, several of which play diverse roles in the probiotic activity of this strain as assessed mainly in preclinical models. Those include the major secreted proteins p40 and p75 that prevent cytokine-induced inflammatory damage, lipoteichoic acid (LTA) that negatively modulate colitis, CpG-rich DNA motifs that dampen allergen-specific IgE, and exopolysaccharides (EPS) that reduce adipogenesis in high-fat-diet fed mice [90]. *L. rhamnosus* GG *spaCBA* encoded sortase-dependent pilin anchored at the surface of the bacteria are important NF as they are involved in the mucus and intestinal epithelium adhesion capacity of the strain. However, they were also shown *in vitro* to act as EM as they contribute to the immunomodulatory capacities of the strain when incubated with monocytes and dentritic cells [105,106], and stimulate cell proliferation that protects against radiologically induced intestinal epithelial damage [107]. Nevertheless, the regulation of expression and functional properties of the pili as well as the other EM during manufacturing of *L. rhamnosus* GG remains largely unexplored.

For example, the regulation of the genes encoding p40 and p75 remains unknown. Additionally, these bioactive molecules are derived from cell wall associated muramidases and are at least partially secreted in the culture supernatant [108], which is usually removed during dried probiotic manufacturing [71], raising doubts about their functional availability in supplement products. Similarly, although the role of specific genes involved in LTA biosynthesis in lactobacilli (including *L. rhamnosus* GG) has been studied [109,110], their regulation by environmental conditions remains largely unknown. For example, Dlt mediated LTA D-alanylation has been long recognized as an important modulator of the host-effects elicited by LTA [111], but we do not know whether *dlt* expression is regulated by environmental conditions in *L. rhamnosus* GG. Notably, it has been demonstrated that the *dlt* gene of *Staphylococcus aureu*s, is regulated by cations levels in the medium (Na+, Mg(2+), Ca(2+)) [112], and in *L. casei* (a close relative of *L. rhamnosus*) is regulated by the presence of charged molecules like antimicrobial peptides [113]. These findings suggest that *dlt* regulation in lactobacilli may be coordinated similarly to what was observed in *S. aureus*, and that the concentrations of positively charged components in the growth medium may affect D-alanylation of LTA. A specific galactose-rich exopolysaccharide in *L. rhamnosus* GG has been previously identified to play a role in the adhesion capacity of the strain [114].

Although regulation of the production of this EPS in *L. rhamnosus* GG has not been studied in detail, recent studies in *L. plantarum* VAL6 indicated that expression of *eps* genes eliciting structural changes of the polysaccharides produced in this species is regulated by pH and sodium chloride induced stress [115], conditions that may occur during industrial growth. Importantly, manufacturing was demonstrated to influence the presence of the SpaCBA pili at *L. rhamnosus* GG's surface and could thus affect the presence and function of this important niche factor and effector molecule in preparations of this strain. It has been shown that centrifugation at 8000g for 30 min was sufficient to break and separate the pili from the surface of the bacteria [100], while a specific type of drying (spray-drying without addition of any protectants, which is not a common manufacturing practice) diminished the adherence capacity of *L. rhamnosus* GG correlated with the disappearance of the SpaCBA pili [116] (Figure 2). It is not known today if the presence of this importance protein structure can be influenced by other types of drying. However, freeze-drying was shown to decrease the adherence capacity of *L. rhamnosus* GG, while it did not exert the same effect on *L. casei* Shirota [117]. Moreover, besides the physical presence or absence of the pili structure in preparations of this strain, the genetic region encoding the SpaCBA pili was shown to be relatively unstable [118], which may also contribute to variations in *in vivo* behavior of *L. rhamnosus* GG isolates originating from different products [119]. Altogether, these lines of evidence indicate that upstream (e.g., fermentation conditions) or downstream processing conditions (e.g., centrifugation, type of drying) can play a role in the presence and bioavailability of NF as well as EM of *L. rhamnosus* GG. *In vivo* demonstration of the impact of those processing induced modifications has not yet been pursued, but we hypothesize that they may have contributed to the different outcomes obtained in clinical trials like those reported by Kalliomäki et al. [97,98] and Kopp et al. [99].

L. plantarum WCSF1 growth phase influences its host immunodulatory capacity

To the best of our knowledge, the only substantiated example demonstrating that production parameters (i.e., growth phase harvesting) can influence the way a probiotic can interact with the human body has been obtained with the well-characterized *L. plantarum* WCFS1 strain. Freeze-dried preparations of heat-killed or live *L. plantarum* WCFS1 were administered to healthy adults. In addition, the live preparations consisted of cells harvested during mid-logarithmic or during the stationary phase of growth. Following consumption of these distinct preparations of the same strain, duodenal tissue biopsies were analysed by array-based transcriptomics, revealing that both live and dead (heat treated) stationary phase harvested bacterial preparations were able to modulate Nfκ-B responses in human duodenum mucosal tissues, which were interpreted to play an important role in the establishment of immune tolerance. Conversely, the bacteria harvested mid-exponentially failed to induce such responses, but modified the expression of human genes involved in immune-suppressing activities such as BCL3, Iκ-B and ADM, as well as several functions involved in cell-cycle and metabolic regulation [120]. As a follow-up, it was found that the *lp_0800* gene, coding for a serine- and threonine-rich surface protein (StsP) that is anchored to the peptidoglycan by sortase was shown to be expressed predominantly during the stationary growth phase, albeit at low levels during growth under laboratory conditions. Notably, previous studies of *L. plantarum* had established that the expression of *lp_0800* was *in situ* induced during the transit through the murine and human intestinal tract, supporting that specific environmental conditions can modulate its expression [121-123]. Importantly, using isogenic *L. plantarum* WCFS1 lacking or overexpressing StsP, it was shown that StsP surface derived peptides obtained by whole-cell trypsin-shaving could strongly inhibit flagellin induced Nfκ-B activation in a CaCo-2-derived reporter cell line. Finally, gel-purified StsP protein derived tryptic peptides potently suppressed NFkB activation, unambiguously pinpointing this activity to peptides derived from this surface protein [89]. These findings demonstrate that the growth phase as well as specific growth conditions (i.e., gut-like conditions) of *L. plantarum* WCFS1 can influence the expression level or bioavailability of the important immunomodulatory StsP, which was proposed to play a prominent role in the clinically observed duodenal transcriptional responses (Figure 2).

Of note, the host responses were determined in the duodenum of the participating volunteers. In fact, upon ingestion, the relatively short transit time to reach the duodenal mucosa likely allows a limited molecular adaptation of the probiotic bacteria. At this moment it is unclear whether similar transcriptome response differences would be observed in the colonic mucosa when applying these distinct *L. plantarum* WCFS1 preparations. One the one hand, the different molecular make-up of the preparations may change during gastrointestinal transit, and on the other hand it is known that *stsP* expression is induced in the intestinal tract. We hence hypothesize that ensuring NF and EM presence and function in probiotic products may be especially relevant when the probiotic is expected to elicit its health benefit in the proximal regions of the intestine.

Overall, the *L. plantarum* WCFS1 example strongly indicates that upstream processing (e.g. harvesting time) can impact probiotic bioactivities *in vivo,* and underlines the importance of harvesting probiotic cells at stationary phase, which is today a common practice in industry. Moreover, it supports that quantification of StsP in preparations of *L. plantarum* WCFS1 could serve as a molecular quality control parameter to complement the traditionally used CFU enumeration.

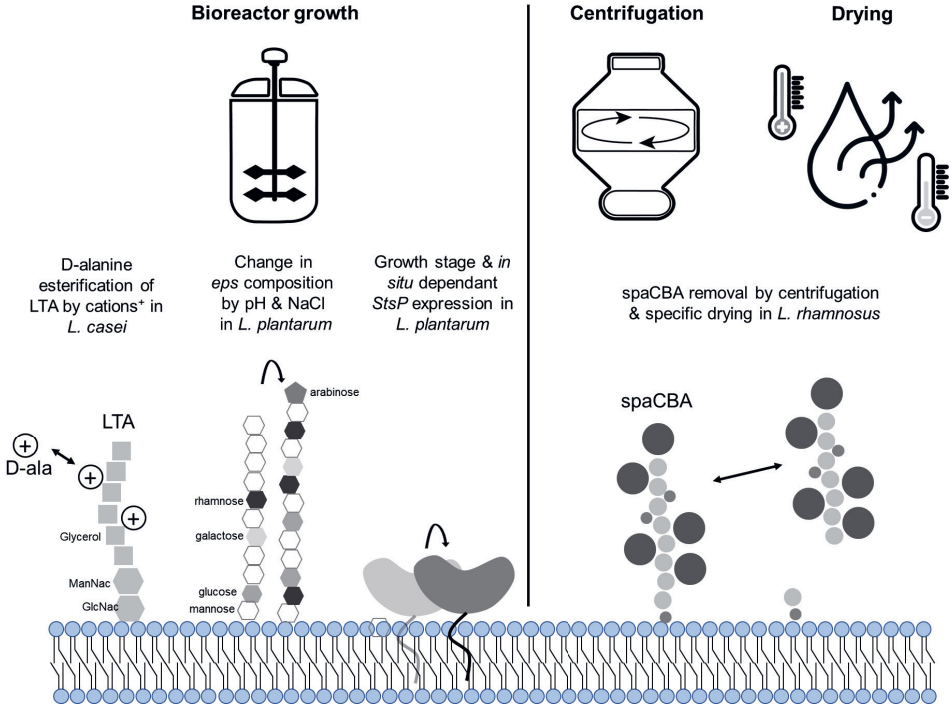


Figure 2 Overview of potential effects of manufacturing on Lactobacillacae NF and EM. *L. rhamnosus* GG SpaCBA pili has been shown to be at least partially removed by centrifugation (8000 g, 30 minutes) or spray-drying. StsP has been shown to be predominantly produced in the stationary growth phase of *L. plantarum* WCFS1, and is regulated in response to intestinal conditions**.** In analogy to what has been shown in *L. casei* and *L. plantarum*, it can be proposed that lipoteichoic acid (LTA) and exopolysaccharides (EPS) structures of *L. rhamnosus* GG can be modified by cations and/or other positively charged compounds.

And what about *Bifidobacterium* ?

Similar to probiotics belonging to the *Lactobacillaceae* family, we propose that manufacturing procedures used for bifidobacteria should be investigated for their possible contribution to the discrepancies observed in clinical trial outcomes [42]. To date, there are only few studies focusing on the effect of manufacturing on bifidobacteria bioactivities and most focused on the potential impact of downstream processing (i.e., drying). Moreover, the available studies did not include an assessment of specific effector molecule presence and bioavailability but were mostly driven by functional assays. For example, Laconelli et al. showed that different drying procedures (air-, freeze-, spray-drying) impacted on the anti-inflammatory properties of *B. bifidum*, as determined by cytokine production profiling in Peripheral Blood Mononuclear Cells (PBMC) following co-incubation with the differently processed bacterial preparations of the same strain. In addition, this study demonstrated that different down-stream process may affect the hydrophobicity of the strain, which is indicative of changes in cell-surface properties [124]. Conversely, similar analyses showed that spray-drying did not alter the immunomodulatory potential of two *B. animalis* subsp. *lactis* strains (INL1 and BB-12), nor did it modify their preventive effect on colitis *in vivo* [125]. Similarly, although freeze-drying was proposed to enhance the adherence capacity of *B. animalis* subsp. *lactis* BB-12, increasing its capacity to outcompete *C. difficile in vitro* [117], these effects were not observed for other strains of *B. animalis* subsp. *lactis* [126]. These studies illustrate the rather limited information concerning the potential impact of manufacturing on bifidobacteria bioactivity, and highlight the contradictory findings described to date on the potential role of downstream processing (e.g., drying) in influencing *Bifidobacterium* probiotic functionality. However, these studies mostly addressed the consequences of different downstream-processing conditions (i.e., drying procedures) on *in vitro* outcomes [127,128], whereas the upstream processing (e.g., fermentation parameters) effects on the expression of NF and/or EM in these bacteria remain to be deciphered.

Metabolic regulation of pili production. Two types of pili have been described in bifidobacteria to act as NF and EM, the sortase dependant pili and the Type IV TAD pili. Sortase dependent pili gene clusters consisting of major (*fimA* or *fimP*) and minor (*fimB* or *fimQ*) subunit structural proteins are widely distributed amongst *Bifidobacterium* species [129]. However, their genetic distribution among strains and species within this genus appears quite disperse. For example, *B. adolescentis* contains five distinct pili encoding gene clusters, while other bifidobacteria, like *B. bifidum*, contains 'only' three of these clusters. Out of the three pili gene clusters found in *B. bifidum* PRL2010, only *pil2* and *pil3* were found to be functional, and *pil1* being disrupted by a frameshift [129-131]. Importantly, and analogous to what was shown for *L. rhamnosus GG*, the sortase dependent pili of *B. bifidum* PRL2010 were demonstrated to play a role in both adhesive and anti-inflammatory properties of the strain using recombinant *L. lactis* harboring the *pil2* or *pil3* gene clusters. While Pil2 was shown to act as a NF and mediated binding to extracellular matrix, Pil3 was also able to modify both *in vitro* and *in vivo* inflammatory responses [131].

Similarly to the sortase dependent pili, type IVb TAD pili are also conserved and widely distributed in both Gram-positive and gram negative bacteria [132], and has been identified in multiple *B. breve* and *B. bifidum* strains [128,133]. It was demonstrated that the Type IVb TAD pili of *B. breve* UCC2003 act as NF, as disruption of the ATPase encoding gene $tadA_{2003}$, which is essential for its assembly, resulted in a decreased capacity of the strain to colonize the mouse intestine [128]. Additionally, using a set of recombinant *B. breve* UCC 2003 strains it was shown that the same pilin structure (and particularly its TadE pilin subunit) could contribute to the maturation of the naïve gut, since it promoted epithelial proliferation both *in vitro* and *in vivo* [134], demonstrating its additional EM role.

Limited information is available about the regulation of production of the various pili that are encoded by bifidobacteria. In *B. bifidum* S17 it was demonstrated that sortase-dependent pili encoded by the *pil2* and *pil3* clusters were higher expressed during exponential growth as compared to the stationary phase of growth [133]. In another strain of the same species, *B. bifidum* PRL2010, the *pil2* and *pil3* clusters were transcribed both during growth in laboratory medium (MRS) as well as in the mouse cecum, whereas transcription of the *pil1* cluster could not be detected under either of these conditions [131]. Culturing of this strain in bovine milk led to activation transcription of the

pil1 cluster genes, indicating that growth (i.e., production) conditions can influence the repertoire of pilin produced by *B. bifidum* PRL2010. This observation was expanded by demonstrating that the other *pil* clusters in this strain were subject to substrate regulation, which is exemplified by the induction of transcription of the *pil2* cluster during growth on fructo-oligosaccharides (FOS) and the induction of the *pil3* cluster during both growth on bovine milk or polysaccharides derived from kefir [129,135]. Analogously, the expression of *pil* gene clusters was also regulated by the carbon source used for growth in *B. adolescentis* 22L, where maltodextrin or cellobiose as substrates for growth resulted in an increase of gene expression (compared to growth on glucose) for *pil3*, *pil4*, and *pil5*, which coincided with increased adhesion of the strain to laminin, fibrinogen and fibronectin, albeit that direct relatedness of these observations remains to be established [130]. Besides the carbon source for growth, also the available nitrogen source in the medium has been reported to control pilin expression. For example, the presence of lysine in the growth medium appeared to be essential for Pil2 and Pil3 production by *B. bifidum* PRL2010 [136].

In both *B. breve* and *B. bifidum*, part of the genes encoding the Type IVb TAD pili were found to be expressed during standard growth conditions in the laboratory [128,133]. However, the transcriptional levels of the Type IVb TAD pilus encoding genes in *B. breve* were strongly induced (25 to 62 fold) when the bacteria were inhabiting the murine intestinal tract. This was further supported by immunogold staining demonstrating that the pili structures could only be observed when the strain was harvested from the murine gut [128]. Even though the Type IVb TAD pili protein presence was not assessed in laboratory-grown (MRS) *B. bifidum* S17, the encoding genes (*tadZ, tadA* and *tadB*) were expressed in a growth phase dependent manner, with higher transcriptional levels in the exponential phase compared to the stationary phase [133].

Although, the specific environmental factors that regulate pili production of specific bifidobacterial pili are quite diverse (e.g., carbon source, nitrogen source, ‘intestinal conditions’, etc.) and/or remain unknown, it is clear from the observations presented above that pilin expression by *Bifidobacterium* probiotics may be modulated by the growth conditions (e.g., substrate) employed during production. In addition, more work deserves to be pursued to decipher the role of downstream processing, as analogous to what has been described for the pili of *L. rhamnosus* GG, we can hypothesis that

the presence of the pili on the cell-surface of the bifidobacteria may be impacted by drying procedures.

Environmental factors regulating serine protease inhibitor production. The serine protease inhibitor (serpin) of pancreatic and neutrophilic elastases was initially described in *B. longum* NCC 2705 [61]. This protein was shown to be conserved in a broad range of bifidobacteria and has been proposed to protect them against host produced proteases, thus providing them with a survival and colonization advantage [61,137]. The serpin's capacity to inhibit the Human Neutrophil Elastase [61] may also be involved in the immunomodulatory capacities of the strain [63] as elastase is released by activated neutrophils at the sites of intestinal inflammation in the gastro-intestinal tract [65]. In line with this role in dampening innate immunity, serpin was demonstrated to play a key role in the anti-inflammatory effect of *B. longum* NCC 2705 in a mouse model of gluten sensitivity [64]. In addition, it was recently reported that the serpin of the NCC 2705 strain prevented enteric nerve activation *in vitro*, which was proposed to potentially play a role in pain reduction in Irritable Bowel Syndrome (IBS) patients [138]. These findings indicate that analogous to the bifidobacterial pili, the role of the serpin is dualistic in the sense that it acts as both NF and EM.

Transcriptional regulation studies of the *B. breve* UCC2003 serpin-encoding gene showed that it involves a protease inducible two-component system encoded directly adjacent to the serpin encoding operon, which was shown to activate serpin production upon exposure of the strain to proteases (e.g., papain) [139]. However, a similar two-component system appears to be absent in *B. longum* subsp. *longum* strains (including NCC 2705), and variable gene-syntenies encountered in the serpin encoded region in different bifidobacterial (sub-)species suggests that serpin regulation may involve (sub-)species specific mechanisms [137]. This notion is further confirmed by our recent study that demonstrated that in *B. longum* subsp. *longum*, serpin production is regulated by the carbohydrate-substrates used for growth, revealing galactose and fructose (or galacto- or fructo- di/oligo-saccharides) as inducing substrates, while the presence of glucose repressed serpin production almost completely [140]. These studies illustrate the diverse environmental factors and regulatory mechanisms involved in controlling serpin production in *Bifidobacterium* (sub-) species, indicating that growth conditions (e.g., substrate) could be tailored to ensure, or even enhance the production and function of this NF and/or EM in the final probiotic preparation.

EPS biosynthesis and potential link to bioactivities. Exopolysaccharides (EPS) are extracellular carbohydrates polymers synthesized by a vast variety of micro-organisms, including Gram-positive bacteria. In bifidobacteria, EPS can be covalently or non-covalently bound to the cell surface (sometimes referred to as capsular polysaccharides; CPS), or can be predominantly secreted. The EPS produced by bifidobacteria are heteropolysaccharides (HePS) that have been reported to vary in molecular weight (between $4.9x10^3$ - $3x10^6$ Da) [141] and monosaccharide composition and linkage [142]. The synthesis of HePS by *Bifidobacterium* strains involves gene clusters (*eps* clusters) and biosynthesis mechanisms that are similar to those described for other microbes, involving a membrane associated synthesis machinery that utilizes cytoplasmic sugar nucleotides as building blocks for the assembly of repeating units that are exported and polymerized to form the HePS [143-145]. Most of the bifidobacterial HePS are predominantly composed of D-galactose and D-glucose but can also contain other monosaccharides like L-rhamnose, D-mannose, L-arabinose, or D-fructose in ratios that can vary between species and likely also between strains of the same species [146,147]. This notion is supported by the fact that *eps* gene clusters are highly variable between different species and strains of bifidobacteria [143-145,148].

Purified EPS produced by *B. longum* BCRC14634 showed an anti-microbial effect on four pathogenic and three food spoiling bacteria [149], which could provide a competitive advantage to the strain in the complex gut ecosystem, supporting its role as NF. In addition, bifidobacterial EPS has also been proposed to affect host responses, suggesting that these molecules may also act as EM in bifidobacteria. Firstly, EPS produced by two *B. animalis* strains (*B. animalis* RH and *B. animalis* subsp. *lactis* BB12) was shown to possess anti-oxidant capacities *in vitro*, [146,147], which could be relevant in order to alleviate intestinal oxidative damages. Then, the EPS produced by different *B. longum* and *B. animalis* strains has been proposed to modulate the immune response of the host based on the role of these molecules in inducing immune cell proliferation and modulating cytokine production in peripheral blood mononuclear cells (PBMCs) [150,151]. Notably, the immunomodulatory effect of EPS observed in this study was shown to be strain specific, which was exemplified by the finding that out of eight strains of *B. longum* and *B. animalis* tested, only the EPS produced by *B. animalis* A1 and *B. longum* NB667 elicited a significant increase of PBMC proliferation [150]. The immunomodulatory capacities of specific bifidobacterial EPS molecules has

been reported to be quite diverse, including reports on the induction of pro-inflammatory profiles *in vitro* [150,152] or *in vivo* [151] but also cases where anti-inflammatory responses were detected [144,149]. However, these studies provided very limited information on the physical and chemical characteristics or the monosaccharide composition of the EPS molecules produced by these different bifidobacterial strains, leaving the relationships of EPS structure and its immunomodulatory function unaddressed. A study by Hidalgo-Cantabrana and colleagues partly elucidated how EPS characteristics might affect its biological activity. In this study they used *B. animalis* subsp. *lactis* A1 and two mutant derivatives (A1dO and A1dOxR) that produce EPS with distinct monosaccharide composition and molecular size characteristics compared to the wild-type strain [153]. Strain A1dOxR harbored a mutation in the *Balat_1410* tyrosine kinase encoding gene [154] and expressed the enzyme dTDP-glucose 4,6-dehydratase that catalyzes the production of dTDP-rhamnose at an elevated level, leading to the production of a high molecular weight, rhamnose-rich EPS. The changes in polymer length and monosaccharide composition in this mutant strain were associated with increased production of the anti-inflammatory cytokine IL-10 by PBMCs exposed to the A1dOxR strain relative to its parental strain [153]. To the best of our knowledge this is one of the few studies where the EPS structure-function relationship is investigated in the context of immunomodulatory capacities, illustrating that there is a large gap in our mechanistic understanding of the postulated role of EPS as EM in bifidobacteria.

Several studies have demonstrated that growth conditions modulate both the level of production as well as the structural properties of EPS produced by bifidobacteria. As an example, the carbon substrate applied during growth influences the expression of *eps* related genes in *B. longum* CRC002 [155], which is potentially influencing the EPS produced by the strain. Indeed, the carbon source used for growth of *B. longum* BB79 affected the level of EPS produced, with lactose leading to the highest level of production when compared to glucose, fructose, or sucrose [156]. Besides the influence of the carbon source, differences in concentration of yeast extract, growth temperature and incubation time modulated the EPS production by *B. animalis* BB12 [157], and the level of dissolved oxygen and CO2 concentrations affected EPS production in *B. longum* JBL05 [158]. Although these studies did not investigate the potential compositional changes in the EPS that was produced, they do highlight that various

growth conditions can influence EPS production. In this context, it should be noted that in *L. rhamnosus* E/N, the carbon source used for growth did not only affect the quantity of EPS produced but also its monosaccharide composition [159]. Taken together, the existing information illustrates the limited and scattered knowledge of the regulatory mechanisms underlying the production of (different) EPS molecules in bifidobacteria. Especially when the role of EPS as a NF or EM is to be further substantiated, better understanding of the regulation of EPS production and composition will be required to reliably investigate the role played by these molecules in *Bifidobacterium* probiotics. Such knowledge would also be required to design manufacturing procedures that aim to improve the presence and abundance of bioactive EPS molecules in *Bifidobacterium* probiotic products, in order to enhance their health benefit reliability as we propose in this review.

Environmental regulation of surface enzymes involved in adhesion. Different cytoplasmic enzymes are also found on the surface of bacterial cells, such as transaldolase, enolase, and DnaK. These surface attached proteins were suggested to act as bifidobacterial NF, based on their *in vitro* demonstrated role in the adhesive properties of *Bifidobacterium* strains. The example proteins mentioned serve typical cytoplasmic functions in glycolysis (transaldolase and enolase) or stress response (DnaK is a chaperonin), but were suggested to be secreted via a yet unknown nonclassical secretion mechanism [160]. This class of surface exposed cytoplasmic proteins is often referred to as moonlighting proteins [161,162] defined by the fact that they can perform two or more physiologically relevant biochemical or biophysical functions [163]. It could be that these proteins become deposited on the cell surface upon lysis of surrounding bacteria that release their cytoplasmic content in the environment. Alternatively, it was recently proposed that in *B. longum* NCC 2705 those type of surface exposed cytoplasmic (moonlighting) proteins could be excreted through the formation of extracellular vesicles [56], of which the formation was shown to occur through membrane bubbling upon peptidoglycan damage in *Bacillus subtilis* [164].

Modulation of the expression of these cytoplasmic proteins could also affect their surface exposure levels. Although such relation has to the best of our knowledge not been investigated in detail, it has been reported that different environmental factor (pH, bile salts) or mild stress conditions [76,165] can increase expression levels of moonlighting proteins. Finally, extracellular vesicles are one of the purported export-

mechanism for adhesive proteins and may be modulated during growth as the presence of yet unidentified gut microbiota derived metabolites (in *in vitro* fecal fermentations) increased their formation [56]. More work is required to understand how these proteins are ending up on the cell surface, and if different manufacturing steps (e.g., steps inducing lysis of cells or growth condition inducing extracellular vesicles formation) might influence this process.

Table 1 : Summary of known bifidobacterial NF and EM, their validation level and related evidence supporting an effect of manufacturing

Protein	**Bioactivity class & validation level**	**Evidence of potential manufacturing Impact**
Sortase dependent pili	Niche factor; *in vitro* [131]	Growth phase dependent transcription in *B. bifidum* [133].
	Effector molecule; *in vitro* & *in vivo* [131]	Carbohydrate regulated transcription in *B. bifidum* [129,135] and *B. adolescentis* [130]
		Lysine presence is necessary for protein production in *B. bifidum* [136]
Type IVb TAD pili	Niche factor; *in vivo* [128]	Growth phase dependent transcription in *B. bifidum* [133]
	Effector molecule; *in vitro* & *in vivo* [134]	
Serpin	Niche factor; in vitro [61]	Carbohydrate substrate controls protein presence in B. longum [140]
	Effector molecule; in vivo [64]	Protease presence controls expression in *B. breve* [137,139]
Moonlighting proteins (transaldolase, enolase, DnaK)	Niche factor; *in vitro* [161,162,166]	Unknown metabolite presence during growth enhance extracellular vehicle production in *B. longum* [56]
EPS	Niche factor; *in vitro* [146,147,149]	Carbohydrate substrate modified transcription [155] and yield in *B. longum* [156]
	Effector molecule; *in vitro* [150,152,167] & *in vivo* [151], but lack structure/fonction relationship	Growth conditions modulate EPS yield in *B. animalis* [157] and *B. longum* [158]

Conclusions & discussion

Discrepancies in clinical trial outcomes stemming from the use of probiotics belonging to the *Lactobacillaceae* family or the *Bifidobacterium* genus has been previously reported. Variations in trial design (population, dose, outcome measurement, etc.) have been advocated to at least in part explain the differences in the outcomes. However, the information summarized in the present review indicates that manufacturing conditions can influence on the presence and/or function of probiotic molecules that play critical roles in the survival of the bacteria in the intestinal tract (NF) and/or their interaction with host cells (EM), which may in turn be an additional cause for the observed variations in clinical outcomes. This review argues that knowledge of probiotic NF and/or EM molecules can provide means to assess the impact of manufacturing conditions on the functionality of the studied strain. Ensuring the presence of validated NF and EM in the final probiotic product could ultimately contribute to improve the consistency of the probiotic's clinical effect.

As mentioned above, ensuring the presence and function of NF and EM during manufacturing could be particularly relevant for upper-gastrointestinal tract mediated health-benefits. An interesting example supporting this hypothesis is the serpin produced by *B. longum* NCC 2705 that was shown to reduce gliadin-induced immune responses in a mouse model, which was proposed to compensate the duodenal serine protease inhibitor decrease observed in active celiac disease. Recently, it was shown that the level of serpin production by this strain can be strongly modulated by the carbon source applied for growth, which allows the production of serpin-rich and serpin-poor probiotic preparations that can subsequently be evaluated *in vivo*. Such approach could establish the relevance of EM presence and function in the probiotic product for health benefits elicited in the proximal region of the intestine. Conversely, the Type IVb TAD pili of *B. breve* UCC 2003 was demonstrated to act as EM in the promotion of colonic epithelial cell proliferation. In this case, it would be interesting to study if the presence or absence of those EM in the initial preparation influences their mediated effect in the distal regions of the intestine, as pili expression was shown to be induced during intestinal transit.

Overall, this review highlights that beyond an improvement of the current clinical trial designs, there is a need to better understand the impact of manufacturing on clinical

efficacy of probiotic products. First, NF and EM molecules have to be established as surrogate markers for probiotic functionality, linking their presence or absence to functional readouts. Work around this topic is relatively advanced for *Lactobacillaceae* (e.g., with the cases of *L. plantarum* WCFS1 and *L. rhamnosus* GG). However, for *Bifidobacterium* probiotics, only a few molecules were identified to act as EM in preclinical animal studies, like the sortase dependant pili of *B. bifidum* PRL2010, the Type IVb TAD pili of *B. breve* UCC 2003 and the serpin of *B. longum* NCC 2705. Overall, additional work, including *in vivo* demonstrations, is needed to identify and validate the molecules driving the host-bifidobacteria interactions. Once NF or EM established as surrogate markers of functionality, implementation of molecular level quality controls (i.e., based on NF and EM), could nicely complement the traditional live cells (E.g. CFU) enumeration in the final probiotic preparation, hence providing more insight on the functional consistency of dried probiotics.

CHAPTER 3

3

Carbohydrate-controlled serine protease inhibitor (serpin) production in *Bifidobacterium longum* subsp. *longum*

S. Duboux [1,2] | M. Golliard [1] | J.A. Muller [1] | G. Bergonzelli [1]|
C. J. Bolten [1] | A. Mercenier [2] | M. Kleerebezem [2]

[1] Nestlé Research, Route du Jorat 57, CH 1000 Lausanne 26, Switzerland
[2] Host-Microbe Interactomics Group, Wageningen University & Research, De Elst 1, 6708WD Wageningen, The Netherlands

This chapter was published as Duboux S, Golliard M. Muller J.A, Bergonzelli G., Bolten C.J., Mercenier A., Kleerebezem M. Carbohydrate-controlled serine protease inhibitor (serpin) production in Bifidobacterium longum subsp. longum. Sci Rep 11, 7236, doi:10.1038/s41598-021-86740-y (2021).

Abstract

The Serine Protease Inhibitor (serpin) protein has been suggested to play a key role in the interaction of bifidobacteria with the host. By inhibiting intestinal serine proteases, it might allow bifidobacteria to reside in specific gut niches. In inflammatory diseases where serine proteases contribute to the innate defense mechanism of the host, serpin may dampen the damaging effects of inflammation. In view of the beneficial roles of this protein, it is important to understand how its production is regulated. Here we demonstrate that *B. longum* NCC 2705 serpin production is tightly regulated by carbohydrates. Galactose and fructose increase the production of this protein while glucose prevents it, suggesting the involvement of catabolite repression. We identified that di- and oligosaccharides containing galactose (GOS) and fructose (FOS) moieties, including the human milk oligosaccharide Lacto-N-tetraose (LNT), are able to activate serpin production. Moreover, we show that the carbohydrate mediated regulation is conserved within *B. longum* subsp. *longum* strains but not in other bifidobacterial taxons harboring the serpin coding gene, highlighting that the serpin regulation circuits are not only species- but also subspecies- specific. Our work demonstrates that environmental conditions can modulate expression of an important effector molecule of *B. longum*, having potential important implications for probiotic manufacturing and supporting the postulated role of serpin in the ability of bifidobacteria to colonize the intestinal tract.

Introduction

Bifidobacterium longum subsp. *longum* is a Gram-positive, high G+C content, anaerobic bacterium that is a member of the phylum Actinobacteria. A variety of bifidobacteria species are natural inhabitants of the human gastro-intestinal tract (GIT). In particular, they are dominantly colonizing the infant gut. Bifidobacteria are believed to be acquired by vertical transmission from the mother and their persistence in the infant gut is associated with their saccharolytic activity toward the abundant host and diet-derived glycans present in the infant gut [168]. Human Milk Oligosaccharides (HMOs) are complex carbohydrates found in human milk where they are the third most-abundant solid component [169]. HMOs, as well as lactose, galacto- [170,171] and fructo-oligosaccharides [172] in infant formula have been shown to drive *Bifidobacterium* enrichment in the infant gut. In addition, bifidobacteria remain present in the human GIT throughout life and various prebiotic dietary supplements (including galacto- and fructo- oligosaccharides) also enrich the bifidobacterial populations in adulthood. According to recent estimates and depending on the geographical location, *B. longum* subsp. *longum* can make up to 20 % of the *Bifidobacterium* community in the intestine, which can constitute up to 4% of the overall microbiota in adults [168]. Strains belonging to the *Bifidobacterium* genus, including those belonging to *B. longum* subsp. *longum*, are widely used as probiotics and their beneficial effects on specific intestinal and extra-intestinal pathologies have been documented [6].

The genome of the strain *B. longum* subsp. *longum* NCC 2705 (hereafter *B. longum* NCC 2705), isolated from infant feces, highlighted its particular adaptation to saccharolytic metabolism relevant in the human gut environment, as illustrated by a large repertoire of genes encoding carbohydrate degrading enzymes and import systems [16,19]. The strain also contains different systems enabling it to resist and evolve in the competitive GIT environment. For example, the genome encodes a bile salt hydrolase and the corresponding efflux system conferring bile acid resistance [54], as well as genes encoding fimbriae that have been shown to drive its *in vitro* adhesion to mucins [57]. This strain is also able to limit growth of the pathogenic *E. coli* O157 *in vivo* through the production of acetate, which depends on the gene encoding the sugar ABC transporter solute-binding protein (*BL0033*) [26]. Moreover, *B. longum* NCC 2705 produces a serine protease inhibitor (serpin) encoded by the *BL0108* gene, which forms covalent products with pancreatic and neutrophil elastases thereby inhibiting

their function [61]. This serpin is proposed to play an important role in the colonization of bifidobacteria by protecting them against host-derived proteases and providing them with a survival advantage in the competitive intestinal environment [61,137]. The serpin's capacity to inhibit the Human Neutrophil Elastase [61] may also be involved in the immunomodulatory capacities of the strain [63] as elastase is released by activated neutrophils at the sites of intestinal inflammation [65]. In line with this role in dampening innate immunity, serpin was demonstrated to play a key role in the anti-inflammatory effect of *B. longum* NCC 2705 in a mouse model of gluten sensitivity [64]. Recently, the serpin of the NCC 2705 strain was reported to prevent enteric nerve activation *in vitro*, which suggest a potential role of this protein in pain reduction in Irritable Bowel Syndrome patients. [138].

In view of the hypothesized colonization and bioactive roles of serpin in bifidobacteria, it is important to understand how its production is regulated. Serpin encoding genes have been identified in a limited number of bifidobacterial species, including *B. breve*, *B. longum* subsp. *longum*, *B. longum* subsp. *infantis*, *B. longum* subsp. *suis* and *B. dentium* [137]. However, querying the genomes of a range of recently described bifidobacterial species [173] identified several additional serpin-encoding genes, suggesting that the function is more widely spread among the members of this genus (Table S 1). A previous study failed to detect serpin production by western blot in MRSc grown *B. longum* NCC 2705 and serpin was thus considered to be induced *in vivo* by unknown factors (F. Arigoni, personal communications). Transcriptional regulation studies of the *B. breve* serpin-encoding gene showed that it involves a protease inducible two-component system located next to the serpin encoding operon [139]. This system is absent from *B. longum* subsp *longum* strains and is partially present in strains of *B. longum* subsp. *infantis*. In all studied species, it was shown that the serpin encoding gene is flanked by genes encoding a Lac-I regulator and a membrane-associated protein of unknown function, although the latter gene is not universally conserved [137]. The observed variations in the gene syntenies around the serpin gene suggest that its regulation may not only differ in different species, but may even be subspecies specific [137]. We have here revealed that different carbohydrates control production of serpin, a protein playing a key role in the bioactivity of *B. longum* subsp. *longum.*

Results

Development of a sandwich ELISA enabling serpin quantification. Polyclonal antibodies obtained by immunizing rabbits with the previously purified *B. longum* NCC 2705 serpin recombinant protein [61] were used to develop a sandwich ELISA (Figure S 1). This method enabled the accurate quantification of the serpin protein in subcellular fractions and crude lysates of *B. longum* NCC 2705, with a limit of quantification (LoQ) of 4 pg/ml of serpin. Moreover, it allowed the detection and quantification of the serpin from *B. breve* ATCC 15700, a protein with only 93% sequence identity compared to the *B. longum* subsp. *longum* serpin [137] (Figure 1).

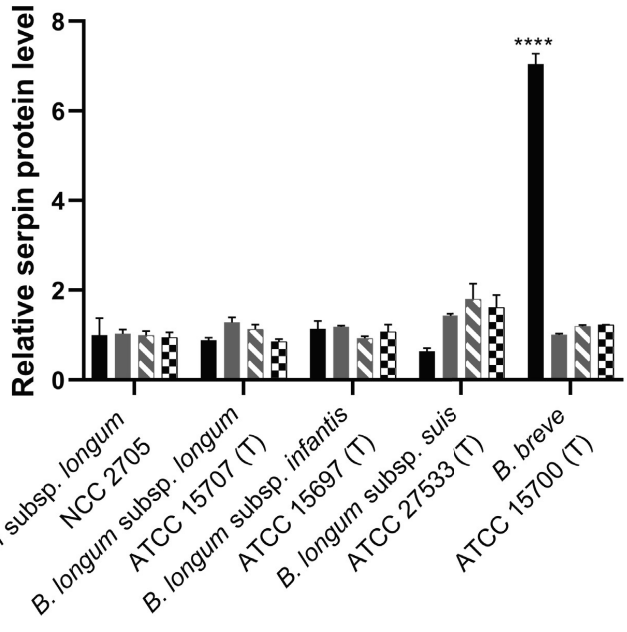


Figure 1: Effect of different proteases (papain [black bars], trypsin [grey bars], chymotrypsin [dashed gray bars] and pig pancreatic elastase [PPE, squared black bars]) on the serpin production levels in bifidobacteria harboring the serpin encoding gene, after 16h of growth. The studied strains are *B. longum* subsp. *longum* NCC 2705, B. *longum* subsp *longum* ATCC 15707 (type strain [T]), *B. longum* subsp. *infantis* ATCC 15697 (T), *B. longum* subsp. *suis* ATCC 27533 (T) and *B. breve* ATCC 15700 (T). Serpin levels are expressed relative to the control level (including standard deviations) measured in MRS without addition of proteases. Statistical difference relative to the control cells are indicated; **** p<0,0001.

Localization of serpin in *B. longum* NCC 2705. Examination of the *B. longum* NCC 2705 BL0108 serpin protein sequence using Interpro-scan [174] predicted that the protein contains a N-terminal transmembrane domain [61]. Analysis by the Signal P [175] package did not predict a signal peptidase cleavage site (Figure S 2), suggesting that the protein would be N-terminally anchored by a transmembrane domain. We investigated serpin localization by the comparative analysis of serpin abundance in cell-associated (cytoplasmic, cell-wall) and the spent culture supernatant (secreted protein) fractions (Figure 2). This analysis revealed that both in the wildtype strain *B. longum* NCC 2705 and in its pMDY25-harboring serpin overexpressing derivative [64], the serpin protein remains associated with the cells and is not released in the environment which in part confirms the predicted localization. The level of serpin produced by the recombinant strain is drastically (> 10,000-fold) higher than the level found in its wildtype counterpart. Importantly, no serpin could be detected in cell-associated or supernatant fractions from strain *B. longum* subsp. *longum* NCC 9035, a genetically engineered serpin null-mutant of NCC 2705 [64] (Figure 2). The minimal background reactivity observed in the cell-wall associated fraction of the NCC 9035, was assumed to be due to minor cross-reactivity of the polyclonal antibody used.

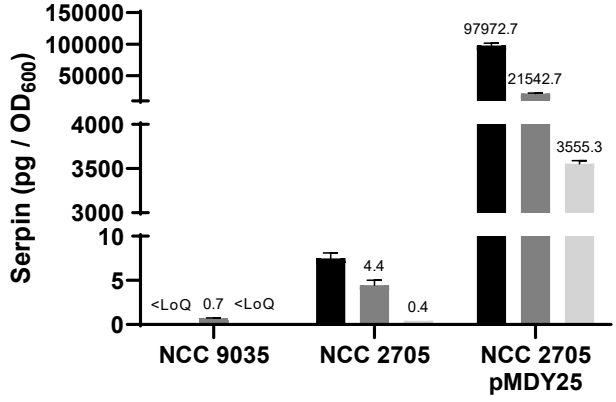


Figure 2: Serpin concentration and localization in extracts of *B. longum* NCC 2705, its serpin knock-out counterpart (NCC 9035) and the constitutively producing (NCC 2705 harboring pMDY25) recombinant strain. Means and standard deviations are depicted. Black bars represent cytoplasmic fractions, dark grey bars represent cell-wall associated fractions and light grey represent supernatant fractions (i.e. excreted protein). Serpin amounts are reported in picograms (pg) per mg of total protein.

Extracellular proteases as inducing factor for serpin production. *B. longum* subsp. *longum* NCC 2705, *B. longum* subsp. *longum* ATCC 15707 (type strain [T]), *B. longum* subsp. *infantis* ATCC 15697 (T), *B. longum* subsp. *suis* ATCC 27533 (T) and *B. breve* ATCC 15700 (T), all belonging to different non-pathogenic species that encode a serpin orthologue closely related to Bl0108 protein (with % of identity of 99.8, 94.9, 92.7 and 93.3 respectively), were grown in presence of different proteases (papain, trypsin, chymotrypsin and pig pancreatic elastase) previously shown to modulate serpin mRNA levels [137,139]. Cell-exposure to papain induced a 7-fold increase in serpin production in *B. breve* ATCC 15700, which is in agreement with the previously reported increased transcription of the serpin gene in this strain [139]. Conversely, none of the studied proteases led to a significant increase in the levels of cell associated serpin in *B. longum* subsp*. longum, B. longum* subsp *infantis* and *B. longum* subsp. *suis* strains, indicating that serpin production is differentially regulated among these bifidobacterial species (Figure 1).

Galactose and fructose driven induction of *B. longum* NCC 2705 serpin is repressed by glucose. To test the capacity of different carbohydrates to induce serpin production in *B. longum* NCC 2705 the strain was grown on carbohydrate-free MRSc medium supplemented with 1% (w/v) of different monosaccharides previously shown to support its growth, i.e., glucose, arabinose, ribose, xylose, galactose and fructose [19]. Growth on the majority of the sugars tested led to a significant increase of serpin levels compared to those measured after growth on glucose (Figure 3A). While only a modest increase in serpin production was observed when the strain was grown on the pentoses arabinose, ribose and xylose, growth on hexoses led to strongly increased levels of serpin, i.e., 16- and 24-fold increase after growth on galactose and fructose, respectively. In these experiments all media were inoculated using a glucose grown preculture which explains the delayed growth observed in some of the conditions. The relatively poor overnight growth observed for *B. longum* NCC 2705 on fructose-media is reflecting the slow-adaptation of the strain to the utilization of this sugar. This was further confirmed by extending the incubation on fructose-media to 48 hours and reaching high OD_{600} values (Figure 3A). Nevertheless, despite the relatively poor growth on fructose media in the initial overnight culture, these growth conditions led to a strongly increased level of serpin (Figure 3A and Figure 3C).

To assess the potential inhibitory effect of glucose, serpin production was next measured after growth of *B. longum* NCC 2705 in media containing different ratios of glucose:galactose (Figure 3B) and glucose:fructose (Figure 3C). In both set of experiments, optical density at 600 nm (OD_{600}) and residual glucose levels in the supernatant were measured after 16h of incubation. We observed that the presence of glucose in the medium prevented the galactose- and fructose-mediated induction, which could be indicative of glucose-mediated catabolite repression. Once the glucose was depleted from the medium (Figure 3B, C), the incubation in galactose and fructose media led to induction of serpin production. Notably, the induction of serpin production in fructose media did not appear to depend on growth on this substrate (Figure 3C).

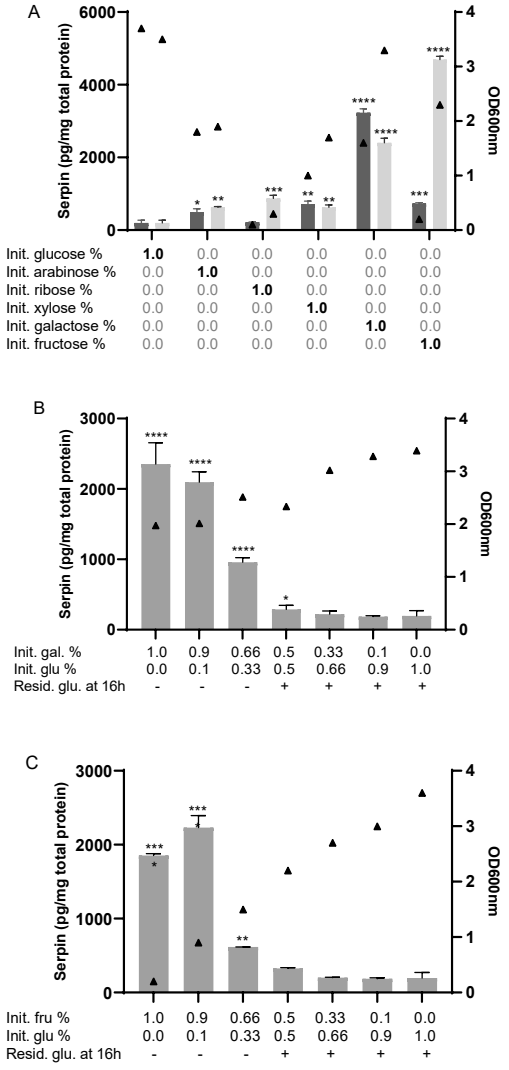


Figure 3: Serpin concentration measured in extracts of *B. longum* NCC 2705 grown for 24 (dark grey) and 48h (light grey) on different monosaccharides (A). Serpin concentration measured in *B. longum* NCC 2705 grown for 16h on different concentrations of glucose and galactose (B) and glucose and fructose (C). Bars represent serpin quantity normalized by total protein content, means and standard deviations are depicted. Initial carbohydrate concentrations are depicted. Presence of residual glucose in the culture supernatant after 16h of incubation is indicated (+/-) (B and C). Triangles represent growth at the end of the incubation, measured by optical density at 600nm. P-values represent statistical difference to the glucose control group. **** p<0.0001; *** p<0.001; ** p<0.01; * p<0.05

Galactose induction mechanism is conserved in *B. longum* subsp. *longum* strains. Different *B. longum* subsp. *longum* strains (Table S 2) were selected to represent the phylogenetic diversity of this subspecies, and these strains were grown on galactose as a sole carbon source for 16h. In these experiments galactose containing media were inoculated with cultures grown in glucose medium. Several strains displayed efficient growth on galactose under these conditions (NCC 2705, CNCM I-2169, NCC 521, ATCC 15708, NCC 552, NCC 293, NCIMB 8809, NCC 305), and in all these cultures a significantly increased level of serpin was detected as compared to their glucose-grown counterparts (Figure 4A). However, galactose did not support growth of all strains within 16h of incubation (ATCC 15707, NCIMB 8810, CNCM I-2171 and DSM 20097), which could reflect strain-specific inability to utilize galactose or a longer lag phase when switching from glucose to galactose. For these strains, a small amount of glucose (0.2%) was added to the galactose-containing (0.8%) medium to allow the production of biomass during the first 16h of incubation, while also including a prolonged period of incubation in galactose containing medium when glucose was depleted. Under these conditions, incubation in galactose containing medium led to increased serpin production (Figure 4B). These results demonstrate that galactose is consistently inducing serpin production in *B. longum* subsp. *longum* strains even in strains that are unable to grow efficiently on this monosaccharide as a sole carbon source.

Similar experiments were performed to evaluate the conservation of the observed serpin regulation in other *Bifidobacterium* species, and/or subspecies of *B. longum*, revealing that galactose-induced serpin production is conserved in *B. longum* subsp. *suis* ATCC 27533 (T) (10,8 fold compared to glucose grown culture) but not in *B. longum* subsp. *infantis* ATCC 15697 (T) or *B. breve* ATCC 15700 (T) (Figure S 2). This demonstrates that differential serpin regulation is not only observed when comparing different *Bifidobacterium* species, but also among different *B. longum* subspecies.

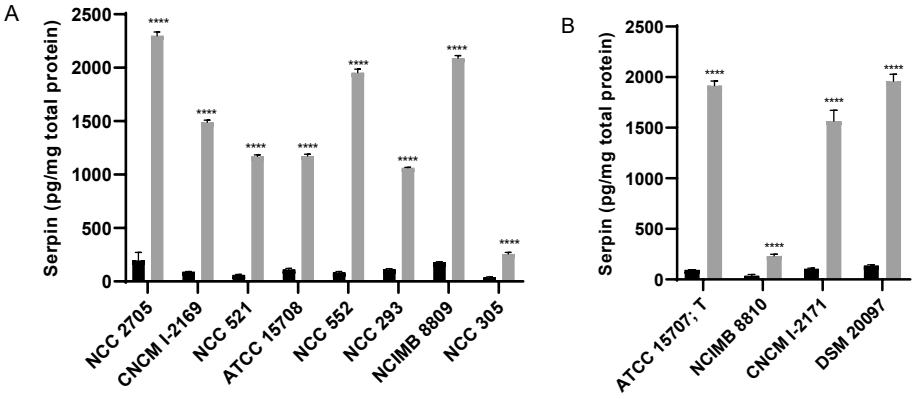


Figure 4: Serpin protein concentration in a set of *B. longum* subsp. *longum* strains grown for 16h on MRSc medium supplemented with 1% glucose (black bars) or 1% galactose (grey bars). Bars represent means and standard deviations are shown. Panel A depicts *B. longum* subsp. *longum* strains able to rapidly grow on galactose as a sole carbon source). Panel B displays *B. longum* subp. *longum* strains not able to rapidly switch to galactose as a sole carbon source. These strains were grown on a medium supplemented with 1% glucose (black bars) or with a mixture of 0.2% glucose and 0.8% galactose (grey bars). P-values represent statistical difference to the glucose control groups. **** $p<0.0001$

Fructose and galactose containing di- and oligosaccharides induce serpin production. *B. longum* NCC 2705 was grown to stationary phase on MRSc medium supplemented with different galactose or fructose containing disaccharides (lactose, melibiose and saccharose), galacto- or fructo-oligosaccharides (GOS and FOS), or human milk oligosaccharides (HMOs). Serpin production was quantified in crude cell extracts derived from the resulting cultures.

Growth on lactose (Gal-β[1→4]-Glc) and melibiose (Gal-α[1→6]-Glc) induced serpin production (respectively 3.7 and 5.2 fold compared to growth on glucose) when supplemented at a concentration of 0.5% (w/v) but not at 1.0 % (w/v). In the latter condition, glucose generated by the hydrolysis of the disaccharides was not depleted at harvesting (16h, Figure 5A). In contrast, saccharose (Glc-α[1→2]-Fru), a fructose-containing disaccharide led to increased levels of serpin irrespective of the concentration used (0.5% or 1.0% supplementation) (Figure 5A).

Growth on different GOS preparations led to a maximal 25-fold induction of serpin in *B. longum* NCC 2705 (Figure 5B). The GOS preparations contain relatively high level of glucose and lactose (King GOS, 28%; Vivinal GOS, 40%; BMOS, 38%). Consequently, also for these substrates the serpin inducing effect was lost at higher supplementation concentrations, correlating with the presence of residual glucose at the time of harvest (16h) (Figure 5B). The FOS preparations tested are purer than GOS and do not contain glucose, explaining why these substrates consistently induced serpin production in *B. longum* NCC 2705 irrespective of the supplementation concentration (Figure 5C).

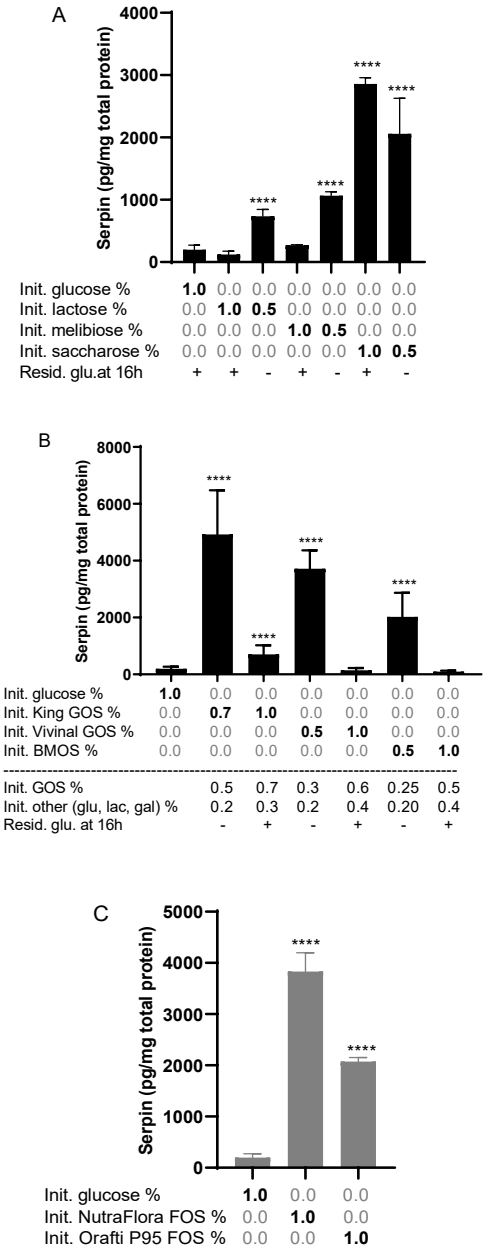


Figure 5: Serpin concentration in extracts of *B. longum* NCC 2705 grown to the stationary phase (16h) on MRSc medium, supplemented with the di-saccharides galactose, melibiose and saccharose (panel A), different galacto-oligosaccharides (GOS) (panel B), or different fructo-oligosaccharides (FOS) (panel C). Bars represent means and standard deviations are depicted. Initial carbohydrate concentrations as well as the presence of residual glucose at harvesting are indicated (A and B). P-values represent statistical difference to the glucose control groups. **** p<0.0001

Finally, *B. longum* NCC 2705 was grown in media containing 0.5% of different HMOs (2'FL, LnNT, LNT, 3'SL and diFL) as sole carbon source. The strain could only grow on Lacto-N-tetraose (LNT), which also led to a 18-fold induction of serpin production as compared to glucose (Figure 6A). LNT was subsequently shown to consistently support the growth and induce serpin production in the different strains representing the phylogenetic diversity of the *B. longum* subsp. *longum* subspecies (Figure 6B).

Taken together these results establish that growth of *B. longum* NCC 2705 (and several other strains of the same subspecies) on galactose and fructose containing di- and oligosaccharides consistently led to increased serpin production. This induction was observed only when the glucose moieties of the galactose containing substrates were completely depleted, which was not the case for the fructose containing carbohydrates.

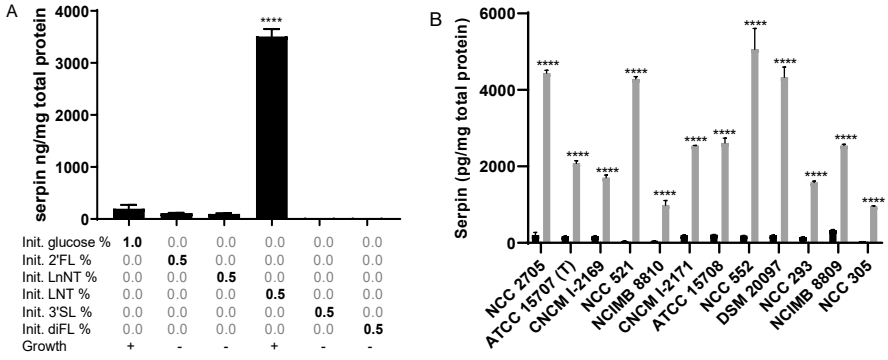


Figure 6: Serpin concentrations in extracts of *B. longum* NCC 2705 grown to the stationary phase (16h) on MRSc medium, supplemented with the human milk oligo-saccharides (2'-O-Fucosyllactose [2'FL], Lacto-N-neotetraose [LNnT], Lacto-N-tetraose [LNT], 3'-O-Sialyllactose [3'SL], Difucosyllactose [DiFL]) (panel A). Different *B. longum* subsp. *longum* strains grown to stationary phase(16h) on MRSc supplemented with 1% glucose (black bars) or 0.5 % LNT (grey bars) (panel B). Bars represent means and standard deviations are depicted. P-values represent statistical difference to the glucose control groups. **** p<0.0001

AraQ involvement in serpin regulation? A MAST search using all transcriptional factors predicted to be implicated in the carbohydrate metabolism of *B. longum* NCC 2705 enabled to identify three LacI-type regulator binding sites in the previously predicted promoter region upstream of the *BI0109* gene [137]. Two binding sites resembling BL0187 and GosR (BL0258) motifs were located far away from the start of the *BL0109* gene (respectively 692 and 1107 bp). In contrast, a motif resembling the predicted AraQ (BL1532) binding site was found at a reasonable distance from the serpin operon (94 bp upstream from the BI0109 start-codon) (Table 1). The resemblance of the potential AraQ binding site is modest when compared to the original BL1532 AraQ binding site but its close localization to the predicted -35 promoter-element supports a role in serpin operon regulation (Figure 7).

Table 1: List of transcription factor binding sites predicted by MAST upfront the promoter region of the *BL0109* gene.

Regulon name	Regulon type	Original sequence (from RegPrecise)	Target sequence upfront of serpin operon	N° of mismatches	Distance from BI0109 start (bp)	p-value obtained from MAST
AraQ (BL1532)	LacI	CCATGTTAACGTTCA CAACT	TCATGGTCACTATG ACAACT	6	-94	3.5E-05
BL0185 (BL0187)	LacI	ATATTGCATCGATGT AAATA	CTATTGCATCGATG CTGATT	5	-692	6.9E-06
GosR (BL0258)	LacI	TTGGTCAACCGGTGT ATCAA	CTGGTCAAGCGTG GTATTAA	5	-1107	4.1E-06

Figure 7: Genetic setup of the serpin operon promoter as previously described by Turroni et al. Inverted repeats (IR 1 and IR 2), -35 and -10 hexamers as well as Ribosome Binding Site (RBS) are depicted in grey. The motif identified using MAST and resembling the previously predicted AraQ motif ahead of the *BL1532* gene is depicted with bold colors.

Conclusions & discussion

In this work, we have developed a sensitive ELISA assay which enabled to quantify levels of serpin present in different bifidobacterial species. Using this assay, we could confirm the high levels of serpin produced by the serpin overexpressing pMDY25-harboring derivative of *B. longum* NCC 2705 [64] as well as the lack of serpin detection in the serpin-null mutant *B. longum* NCC 9035 previously used to decipher the role of serpin in a mouse model of celiac disease [64]. Serpin localization experiments using *B. longum* NCC 2705 showed that the protein is retained in the cellular biomass and not secreted to the medium, which confirmed the predicted N-terminal anchoring of the protein in the cell membrane in combination with the lack of a recognizable signal peptidase cleavage site. However, serpin release in the medium was more prominently detected in the culture of the serpin-overproducing *B. longum* NCC 2705 that harbors pMDY25, which is probably reflecting hampered protein biogenesis as a consequence of the very high serpin production in this strain. Nevertheless, the results establish that to assess regulation of serpin production levels by culture conditions, the protein is best quantified by evaluating its specific-level in the cellular biomass and can ignore the very minor amounts that are released into the medium.

Serpin encoding genes have been found to be conserved among a subset of bifidobacterial species [137]. Previous studies have reported that its transcription in *B. breve* is induced by treatment of the cells with proteases, in particular papain, involving a two-component regulatory system, which is not encoded in *B. longum* [137,139]. The observation in *B. breve* is in apparent agreement with the proposed role of serpin in protection of the bacterial cell against the detrimental impact of environmental proteases. Our results confirmed that serpin production is induced by papain treatment in *B. breve* and established that a similar treatment with papain or other proteases failed to induce it in various *B. longum* subspecies (*longum*, *infantis*, and *suis*). Notably, we could establish that serpin production in *B. longum* subsp. *longum* and *B. longum* subsp. *suis* is controlled by the carbon source in the culture medium, identifying galactose and fructose as potent inducers. Furthermore, our results suggest that import of these carbohydrates in the cell is essential for their effect on serpin production, even though their metabolization appears not required (Figure 3C, Figure 4B). None of the carbon sources tested was able to modulate serpin levels in *B. longum* subsp. *infantis.* This highlights that diverse mechanisms regulate production

of serpin in different bifidobacterial species (e.g. *B. breve* versus *B. longum*), but also in different subspecies. The regulation of serpin production by carbon sources in specific *B. longum* subspecies does not have an obvious relation with the anti-protease function of serpin. This may suggest that an additional functional role of this protein remains to be deciphered.

The data we present suggest the involvement of catabolite repression in the control of serpin production. Growth of *B. longum* on the pentoses arabinose, ribose and xylose led to a modest but statistically significant increase of serpin production compared to glucose, which may be due to relieved glucose catabolite repression. This is in sharp contrast to the prominent serpin-inducing capacity of the hexoses galactose and fructose. Even though both pentoses and hexoses are metabolized through the so called “bifid-shunt”, their metabolism is markedly different; pentoses are entering the energy generating phosphoketolase-dependent pathway as glyceraldehyde-3-phosphate, while hexoses are converted by phosphoglucomutase to glucose-6-phosphate and fructose-6-phosphate [46]. Metabolic intermediates of the “bifid-shunt” may play an important role in the regulation of serpin and would deserve to be further studied in this context.

Although the precise molecular mechanisms involved in serpin regulation in B. longum NCC 2705 remains to be elucidated, we could demonstrate that galactose- and fructose-mediated serpin induction is prevented by presence of glucose in the environment. In B. longum NCC 2705, the glucose/mannose transporter protein (encoded by BL1631; glcP) is involved in galactose import [19]. This protein was demonstrated to have the highest specificity for glucose followed by mannose and galactose [17], which could explain why glucose depletion is required to enable the subsequent import of galactose that once internalized could play a role in the intracellular activation mechanism of serpin production. In the same strain, an ATP-binding cassette transporter encoded by the operon BI0033-0036 (fruEKFG) mediates fructose import. Previous studies have shown that the synthesis of the substrate binding protein (FruE) is strongly suppressed by the presence of glucose [43], suggesting that the fru operon is suppressed by the presence of glucose in the medium. Thereby, the presence of glucose in the medium could inhibit fructose-import, and assuming that internalized fructose (analogous to internalized galactose) could play a role in in the intracellular serpin production-activating regulatory mechanism

provides a rational explanation of glucose suppression of fructose-mediated activation of serpin production (Figure 8).

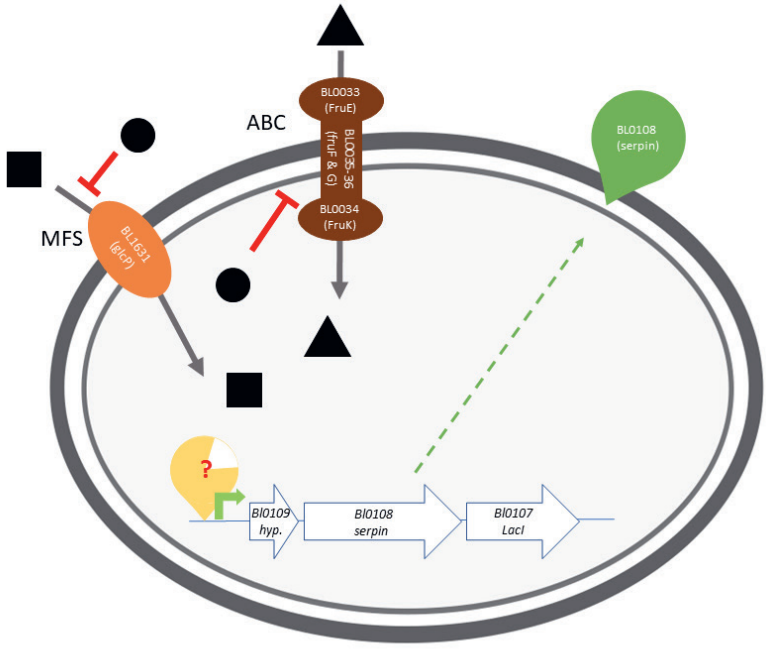


Figure 8: Proposed schematic representation of serpin induction by galactose or fructose and its repression by glucose. Presence of glucose (black circles) inhibits transport of galactose (black squares), as it is known to have a higher affinity to the Major Facilitator Superfamily (MFS) transport protein GlcP (BL1631). As well, presence of glucose in the cell has been previously shown to inhibit the production of the ATP-binding cassette transporters (ABC) encoded by the genes located in the operon *BL0033-BL0036* and responsible of fructose (black triangles) import. The exact regulatory genetic mechanism enabling serpin production remains to be elucidated.

These results resemble typical manifestations of glucose mediated catabolite repression, which is well described in many bacteria [176] and for which a few examples have been described in bifidobacteria[17,177,178]. In most Gram-positive bacteria, catabolite repression is controlled by the complex formed by the histidine phosphocarrier protein (HPr) and catabolite control protein A (CcpA), a LacI-type regulator that can bind to specific DNA motifs (so-called *cre*-elements) and modify transcription of downstream genes [179]. At present, the closest relatives of the *Bacillus subtilis* CcpA in *B. longum* NCC 2705 are annotated as LacI-type regulators, and their homology with CcpA is quite low (24-32% identity). Bifidobacteria encode a large number of LacI-type regulators known to commonly regulate carbohydrate metabolism. For example, *B. longum* NCC 2705 harbors 16 genes encoding LacI-regulators according to the RegPrecise database [180]. Most of these LacI-type regulators are foreseen to act locally by regulating genes located in their vicinity. This is exemplified by the LacI encoding gene *BL0107* that is genetically linked to the serpin encoding gene and is predicted to regulate the closely located sucrose utilization operon that encodes a sucrose permease (BL0106) and a sucrase (BL0105). Recently, two LacI-type transcriptional regulators, AraQ and MalR1, were demonstrated to control a large set of genes spread over the genome of *B. breve* UCC 2003 including genes involved in the ”bifid-shunt” as well as transcription factors, and uptake and metabolism of various carbohydrates [181,182]. Using the online Motif Alignment & Search Tool (MAST) [183], we did not detect MalR1 binding sites, but did identify an possible AraQ binding motif in the promoter region upstream of the serpin-encoding operon (i.e., upstream of *BL0109*). The resemblance of the AraQ cis-element and the sequence found in the serpin promoter region is modest, but in our opinion warrants further experimental investigation to investigate the proposed involvement of AraQ in serpin regulation in *B. longum* NCC2705 and other members of the same (sub-)species.

The galactose containing di- and oligosaccharides (i.e. lactose, melibiose and galacto-oligosaccharides [GOS]) as well as the fructose containing di- and oligosaccharides (i.e. saccharose and fructo-oligosaccharides [FOS]) induce serpin production. Importantly, the presence of glucose in the oligosaccharides prevented galactose induction unless the glucose moiety was completely consumed. GOS are produced using enzymatic synthesis from lactose [184,185] and usually contain high levels of free

lactose and glucose that we demonstrated to suppress serpin production. Conversely, FOS are commonly produced by hydrolysis of long chain inulin and therefore have a very low glucose content [185]. Intriguingly, glucose repression was not observed when cells were grown on saccharose, although this disaccharide is composed of a glucose and a fructose moiety. In contrast to lactose, hydrolysis of sucrose by sucrose phosphorylase (*BL0536*; EC 2.4.1.7) results in the formation of fructose and glucose-1-phosphate in the cell. Intriguingly, specific regulatory roles have been suggested for intracellular glucose-1-phosphate, including its proposed role in the regulation of virulence genes in *Listeria monocytogenes* [186].

Lacto-N-tetraose (LNT) contains two galactose, one glucose and one N-Acetyl-galactosamine residues [169], and hence was tested for its capacity to induce serpin. At an initial concentration of 0.5 % in the culture medium, LNT supported growth and enabled production of serpin in all tested *B. longum* subsp. *longum* strains, indicating a high conservation of the induction mechanism. LNT is likely imported as a tetra saccharide, as the *B. longum* subsp. *longum* strains lacks the gene encoding an extracellular lacto-N-biosidase that would be required for this reaction [187]. Assuming that LNT is imported as a tetra saccharide into the cell, the enzymes involved in its intracellular hydrolysis remain to be identified. However, removal of the terminal glucose moiety may be catalyzed by β-galactosidase. The proposed glucose-mediated catabolite repression may thereby also be transiently activated during LNT metabolization but may have remained unobserved in our experiments due to the low concentration of LNT (0.5%) used. The remaining hydrolysis product (Galβ1–3GlcNAcβ1–3Galβ1) is likely cleaved by the lacto-N-biose phosphorylase (encoded by the gene *BL1641*) into 2 molecules of galactose and one N-acetyl-galactosamine (39). Although we deciphered the role of glucose and galactose in serpin regulation, a possible involvement of N-acetyl-galactosamine remains to be clarified.

Today, even if FOS and GOS are not the oligosaccharides present in human milk, they are commonly used in infant formula as an affordable alternative to HMOs (including lacto-N-tetraose (LNT) [188]) with the purpose to drive *Bifidobacterium* enrichment in the infant gut [170-172]. Our work suggests that these oligosaccharides could have an impact that goes beyond bifidobacterial growth stimulation, by modifying the *in situ* expression of a specific molecule (i.e., serpin), which has been proposed to play a role in colonization and immunomodulatory properties of *B .longum* subsp. *longum* strains.

Materials & methods

Strains and growth conditions. All strains used in this study (Table S 2) were obtained from the Nestle Culture Collection (NCC) (Nestlé Research, Lausanne, Switzerland). Growth was routinely performed using de Man, de Rogosa, Sharpe medium (BD Difco, Franklin Lakes, USA) supplemented with 0.05 % of cysteine (MRSc), at 37°C in anaerobiosis without agitation. Supplementation of the media with 200 µg/ml of spectinomycin was applied for the serpin overexpressing recombinant *B. longum* NCC 2705 that harbors pMDY25 [64]. Where relevant, protease addition rates in the medium were the same than previously described [137], namely 0.5 mg/ml papain, 0.1 mg/ml trypsin, 0.16 mg/ml chymotrypsin (Sigma-Aldrich Chemie GmbH, Buchs, Switzerland) and 1 mg/ml porcine pancreatic elastase (PPE; MP Biomedicals SARL, Illkirch-Graffenstaden, France). Carbohydrate stock solutions were all prepared at 100 g/L in water and filtered sterilized (glucose, arabinose, ribose, xylose, galactose, fructose, lactose, melibiose, saccharose; Sigma-Aldrich Chemie GmbH), galacto-oligosaccharide (GOS) (King GOS [King-Prebiotics, Yunfu City, China]; Vivinal GOS [FrieslandCampina DOMO Amersfoort, The Netherlands]); bovine milk oligosaccharides BMOS [Société des Produits Nestlé, Vevey, Switzerland [171]], fructo-oligosaccharides (FOS) (NutraFlora FOS [Ingredion Korea Inc, Gyunggi-do, Korea]; Orafti P95 [BENEO GmbH, Mannheim, Germany]) and human milk oligosaccharides (HMOs) (2'-O-Fucosyllactose (2'FL), Lacto-N-tetraose (LNT), Lacto-N-neotetraose (LNnT), 3'-O-Sialyllactose (3'SL), Difucoxyllactose (DiFL) [Glycom A/S, Lyngby, Denmark]). These carbohydrates were added to a MRSc-based medium containing no carbohydrates (10 g/L bacto proteose peptone, 3,5 g/L bacto yeast extract, 1 g/L Tween 80 [Chemie Brunschwig, Basel, Switzerland]; 2 g/L di-ammonium hydrogen citrate, 5 g/L sodium acetate, 0.1 g/L magnesium sulphate, 0.05 g/L manganese sulfate, 2 g/L di-sodium phosphate, 0.5 g/L cysteine [Sigma-Aldrich Chemie GmbH]). Two (2) % rate inoculum from an MRSc overnight culture was applied for all growth experiments, which were performed in a BioLector microbioreactor system (m2p-labs GmbH, Baesweiler, Germany), using 48 flowerplate inserted in an anaerobic chamber for 16 to 48h (2ml volume per well, agitation at 600 rpm, CO_2 atmosphere, 37°C). Growth was followed over time by continuous measurement of the scattered light at 620 nm and levels of residual glucose in the culture were determined using the MQuant kit (Sigma-Aldrich Chemie GmbH) according to the manufacturer's protocol.

Subcellular fractions preparation, crude protein extraction and total protein quantification. To determine the subcellular localization of the serpin, cells were grown to stationary phase (16h) and a cell amount corresponding to 12 units of OD600 was harvested by centrifugation (3500 g, 2 min, 4°C). The supernatant was filtered through a 0.2 µm pore size filter (Sigma-Aldrich Chemie GmbH), and the filtrate was precipitated with an equal volume of ice-cold acetone and placed on ice for 20 minutes prior to centrifugation (12,000 g, 10 min, 4°C). The obtained pellet was dissolved with 960 µl of 50mM NaOH (i.e., supernatant fraction). The bacterial pellet was washed once with one volume of ice-cold PBS and resuspended in 960 µl of PBS containing 0.1 mg/ml DNAse I, 20 µg/ml Rnase A and SigmaFast Protease Inhibitor (Sigma-Aldrich Chemie GmbH; 1 tablet per 100 ml of solution). The obtained solution was incubated at 37°C for 30 minutes and was subsequently lysed by bead-beating using a FastPrep-24 (MP Biomedicals SARL), 3 times for 1 minute at 4 m/s, with 2 minutes cooling on ice in between. The obtained crude cell lysate was separated in cytoplasmic (soluble) and crude cell wall (pellet) fractions by centrifugation (12,000 g, 10 min, 4°C). The cell wall fraction pellet was further rinsed twice and resuspended in 960 µl of PBS containing proteinase inhibitor, which was tested to not interfere with the developed ELISA. All fractions were analysed for serpin content (see below) and obtained values were normalized based on culture optical density, measured at 600 nm (OD_{600}).

For all other analyses, stationary phase bacterial cultures (after 16h of growth, unless stated otherwise) were harvested by centrifugation (3500 g, 2 min, 4°C). In order to quantify residual glucose content, supernatants were filter sterilized using a 0.45 µm filter and stored at -20°C. Bacterial pellets were washed with one volume of Dulbecco's Phosphate Buffer (PBS; Sigma-Aldrich Chemie GmbH) and resuspended in 600 µl of PBS containing Halt Protease Inhibitor (Sigma-Aldrich Chemie GmbH). Bacteria were subsequently lysed by bead-beating using a FastPrep-24 (see above; MP Biomedicals SARL). Crude lysates were used as such, containing both soluble and non-soluble fractions. Total protein content was determined using the Pierce BCA protein Assay kit (Thermo Fisher Scientific AG, Basel, Switzerland).

Serpin protein quantification by sandwich ELISA. All anti-serpin rabbit polyclonal antibodies used in this work were obtained from Proteogenix (Schiltigheim, France) and using a purified *B. longum* NCC 2705 recombinant serpin produced in *Escherichia coli* [61]. For the sandwich ELISA, 96 well plates (Nunc MaxiSorp [Thermo Fisher

Scientific AG]) were used. Coating was performed for 16h at 4°C using 100 µl per well of a 0.2 M sodium carbonate/bicarbonate solution at pH 9.4 containing 250 ng/ml of primary anti-serpin antibodies. In between each step, plates were washed 3 times using 300 µl of wash buffer (WB; PBS to which 0.05% of Tween 20 was added). Blocking was performed for 1h at room temperature using 300 µl of blocking buffer (BB; WB to which 2% of Bovine Serum Albumin [BSA; Sigma-Aldrich Chemie GmbH] was added). The standard (purified recombinant serpin [61]) and the different samples were diluted in a final volume of 100 µl of BB and incubated for 2h at room temperature. Secondary biotinylated anti-serpin antibody (250 ng/ml in 100 µl of WB containing 0.2% BSA) was added and incubated for 1h at room temperature, followed by the addition of 100 µl of the enzyme conjugate (HRP Pierce, Thermo Fisher Scientific AG) (10 µg/ml in WB containing 0.2% BSA) and continuing incubation for 1h at room temperature. Subsequently, repeated (6) washes using 300 µl of WB were performed to remove background reactivity. Detection was performed by adding 100 µl per well of HRP substrate (1-step Ultra TMB ELISA, ThermoFisher Scientific AG) and 15 minutes incubation at room temperature. The reaction was terminated by the addition of 100 µl of 2M sulfuric acid and final optical density at 450 nm was measured in a Varioskan spectrophotometer (Thermo Fisher Scientific AG).

Standard curves were obtained following the sigmoidal 4 parameter logistic regression (4PL) method and concentrations of serpin in individual samples were calculated using the following formula:

$$y = \frac{A - D}{\left(1 + \left(\frac{x}{C}\right)^{B}\right) + D}$$

A: Minimal asymptote (y); B: Slope; C: Inflection point (x); D: Maximal asymptote (y)

Statistical analysis. All data obtained for *B. longum* NCC 2705 grown in either MRSc or MRSc + 1% glucose were pooled - (n=36) and were analysed for normality using D'Agostino and Pearson algorithm (data not shown). This data pool was used as control group in every analysis. Data were log transformed to compensate for potential heteroscedasticity and one-way ANOVA was performed on every dataset. P-values were obtained by performing a Turkey's multiple comparison with an alpha level of 0.05.

Bioinformatic analysis. All gene and protein sequences of *B. longum* NCC 2705 referred to in this work were obtained from the National Center for Biotechnology Information (NCBI; NC_004307.2). Operon prediction was obtained from the Prokaryotic Operon Database (ProOpDB) [189]. Functional annotation of the *B. longum* NCC 2705 serpin protein encoded by the *BL0108* gene was performed using Interpro-scan [174] and signal peptide prediction was performed using Signal P [175].

To shed light on the genetic mechanism involved in regulation of serpin production, all bindings sites of transcriptional factors known to be implicated in carbohydrate metabolism regulation were retrieved from the RegPrecise database [180]. These sequences were used as input in the "Motif Alignment and Search Tool" (MAST) [183], to identify resembling motifs in the region upstream of the *Bl0109* gene (genome positions 132852 to 134216; reverse sequenced). Obtained candidate motif-homologues were subsequently mapped relative to the serpin promoter elements (i.e., the -35 and -10 regions) that were previously described by Turroni et al. [137].

Supplementary data

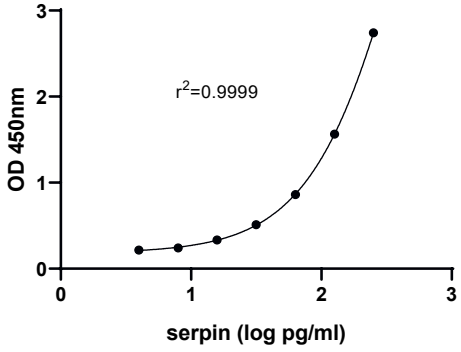


Figure S 1: Standard curve of the developed sandwich ELISA obtained using pure recombinant serpin protein. Nonlinear regression is obtained using the sigmoidal 4 parameter logistic regression (4PL) method.

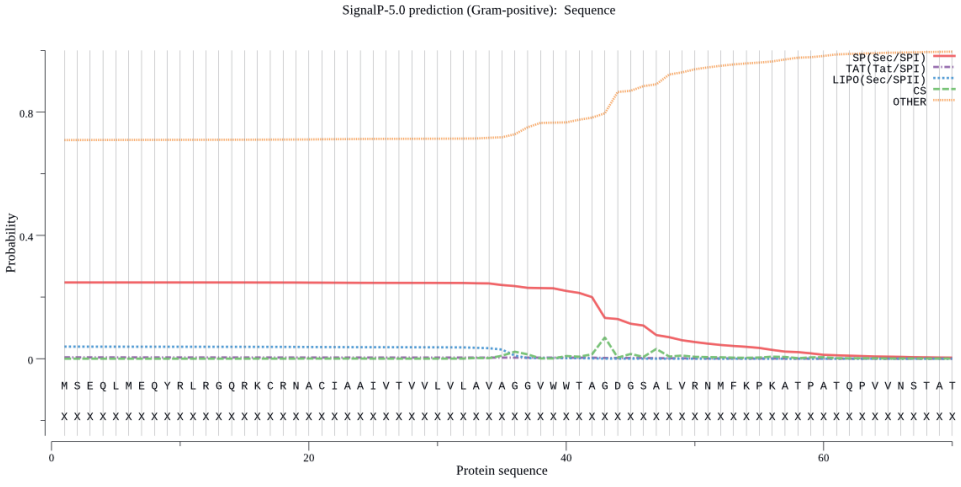


Figure S 2: BL0108 signal peptide & cleavage site prediction using Signal P 5.0. Graphical representation shows that the protein has a low probability of being cleaved (SP; TAT, LIPO, CS), and a high probability of being non-cleaved.

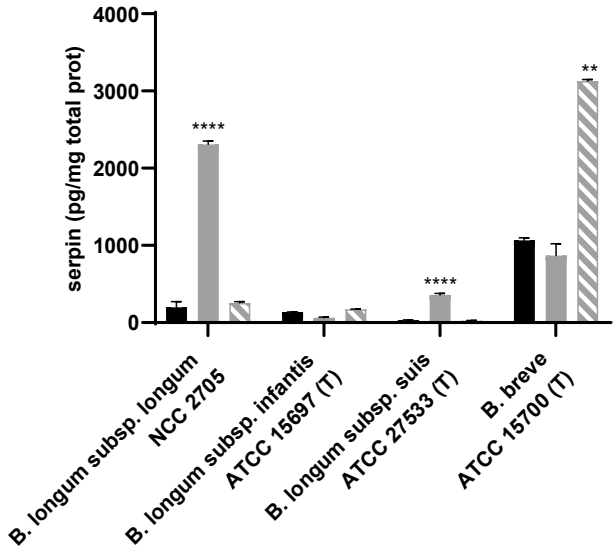


Figure S 3: Serpin levels in extracts of different *Bifidobacterium* species (harboring the serpin encoding genes) grown until stationary phase (16h) on 1% glucose (black bars), 1% galactose (grey bars)) or 1% glucose in presence of 0.5 mg/ml of papain (dashed grey bars). Bars represent means and standard deviations are depicted. P-values represent statistical difference to each glucose control group. **** $p<0.0001$; ** $p<0.01$; * $p<0.05$

Table S 1: Distribution of Bifidobacterium species harboring BL0108 serpin homologous proteins identified in RefSeq Bifidobacterial genomes by BLASTP. Unclassified genomes were excluded from the list.

Species level taxonomy	Number of hits	Number of organisms	Best score
Bifidobacterium longum	97	1	962
Bifidobacterium breve	32	1	904
Bifidobacterium scardovii	2	1	787
Bifidobacterium callitrichidarum	1	1	566
Bifidobacterium scaligerum	1	1	537
Bifidobacterium myosotis	2	1	535
Bifidobacterium felsineum	1	1	531
Bifidobacterium imperatoris	1	1	529
Bifidobacterium saguini	1	1	523
Bifidobacterium goeldii	1	1	481
Bifidobacterium stellenboschense	1	1	402
Bifidobacterium lemurum	1	1	389
Bifidobacterium biavatii	1	1	360
Bifidobacterium dentium	6	1	345
Bifidobacterium moukalabense	5	1	302
Bifidobacterium pseudolongum	27	1	291
Bifidobacterium adolescentis	2	1	252
Bifidobacterium angulatum	1	1	245
Bifidobacterium choerinum	3	1	222
Bifidobacterium vansinderenii	1	1	220
Bifidobacterium callimiconis	1	1	216
Bifidobacterium simiarum	1	1	214
Bifidobacterium criceti	2	1	203

Table S 2: List of strains used in this work.

Strain	Taxonomy	Origin	Ref
NCC 2705 (CNCM I-2618)	*B. longum subsp. longum*	Infant isolate	
NCC 9035	*B. longum subsp. longum*	Bl0108 serpin knock-out derivative of NCC 2705	McCarville et al. 2017
NCC 2705 with pMDY25	*B. longum subsp. longum*	pMDY25 harboring derivative of NCC 2705, Bl0108 serpin constitutive over expression strain	McCarville et al. 2017
ATCC 15707 (T)	*B. longum subsp. longum*	ATCC (typestrain)	
ATCC 15697 (T)	*B. longum subsp. Infantis*	ATCC (typestrain)	
ATCC 27533 (T)	*B. longum subsp. suis*	ATCC (typestrain)	
ATCC 15700 (T)	*B. breve*	ATCC (typestrain)	
CNCM I-2169	*B. longum subsp. longum*	Infant isolate	
NCC 521	*B. longum subsp. longum*	Adult isolate	
ATCC 15708	*B. longum subsp. longum*	ATCC	
NCC 552	*B. longum subsp. longum*	Adult isolate	
NCC 293	*B. longum subsp. longum*	Adult isolate	
NCC 305	*B. longum subsp. longum*	Infant isolate	
NCIMB 8809	*B. longum subsp. longum*	NCIMB	
CNCM I-2170	*B. longum subsp. longum*	Infant isolate	
NCIMB 8810	*B. longum subsp. longum*	NCIMB	
DSM 20097	*B. longum subsp. longum*	DSMZ	

CHAPTER 4

The implication of pH regulation strategies during growth on *Bifidobacterium longum* NCC 2705 serpin production

S. Duboux [1,2] | C. Zandbergen [1,2] | B. Bogicevic [1] | J.A. Muller [1] | A. Mercenier [2] | M. Kleerebezem [2]

[1] Nestlé Research, Route du Jorat 57, CH 1000 Lausanne 26, Switzerland
[2] Host-Microbe Interactomics Group, Wageningen University & Research, De Elst 1, 6708WD Wageningen, The Netherlands

Abstract

We have previously suggested that conditions under which probiotic strains are grown may influence their functional molecular make-up. In this study, we assessed if different pH regulation strategies during growth of *B. longum* NCC 2705 affect the production of serpin, a key bifidobacterial effector molecule. We demonstrated that pH regulation not only increased the final biomass yield, but also affected the obtained levels of serpin. Maintaining a pH of 6 throughout the entire fermentation was key to increase serpin production. In addition, we showed that this pH-mediated induction is controlled by the depletion of glucose during the exponential growth phase. Furthermore, our study revealed a counterintuitive finding, as we observed that controlling the pH over the whole process induces a decrease in cell-permeability and improves the low-pH resistance of the strain. Our study thereby demonstrates that in NCC 2705, different manufacturing conditions aiming at maximizing the final yield also likely influence the strain physiology. This highlights the need to better understand how production process conditions impact on strain properties beyond the traditionally used end point, i.e., colony forming units (CFU). We overall suggest that pH regimes and related carbohydrate consumption patterns should be studied with care in bifidobacterial fermentations to better understand their potential influence on the functional molecular make-up of the strain.

Introduction

Bifidobacterium longum subsp. *longum* (*B. longum*) NCC 2705 strain, isolated from infant feces, was the first bifidobacterial strain to be fully sequenced [16]. Analysis of its genome content reflected its adaptation to the lower human gastro-intestinal tract (GIT). It harbors a large number of genes encoding proteins specialized in transport [19] and catabolism of a variety of carbohydrates, including non-digestible plant polymers or host-derived glycans [16]. Moreover, the strain prefers lactose over glucose, which reflects its adaptation to human milk [17]. Besides these metabolic adaptations, NCC 2705 also encodes several mechanisms supporting its survival and establishment in the GIT, such as surface fimbriae proteins [57] or bile salt hydrolase and bile efflux-transporter [54]. *B. longum* NCC 2705 was also shown to encode a cell-surface exposed serine protease inhibitor (serpin) [61]. The serpin encoded by NCC 2705 was proposed to protect the strain from the stress that may be imposed by host serine proteases produced in the digestive tract [61]. This protein was also shown to play a key role in the immunoprotective capacity of the strain in a preclinical murine model of celiac disease where its isogenic serpin knock-out derivative failed to elicit protection against the negative effects of gliadin exposure [64].

The conditions under which probiotic strains are grown may influence their functional molecular make-up [84]. For example, it was suggested that surface exposed molecules of *Lactobacillaceae* strains may be influenced by the presence of charged molecules in the medium. In *L. casei*, the *dlt* dependent D-alanylation of lipotechoic acid (LTA) was shown to be regulated by the presence of charged molecules in the growth medium [113], similarly to what has been observed in *S. aureus* where cations (Na+, Mg(2+), Ca(2+)) levels in the medium regulated *dlt* expression [112]. Furthermore, expression of *eps* genes eliciting structural changes of the polysaccharides produced by *L. plantarum* VAL6 is regulated by acid (low pH) and sodium chloride induced stress [115]. Overall, these observations highlight that the presence of charged molecules in the medium can influence the surface properties of *Lactobacillaceae* strains and suggest that pH regulation during the fermentation might play a role in the molecular make-up of those strains.

While *Lactobacillaceae* are considered resistant to relatively low pH [190], acid tolerance in bifidobacteria is usually considered to be weaker with the exception of

Bifidobacterium animalis [191] that survives acidic pH better than other bifidobacteria species. Several authors have suggested that application of pH stress on bifidobacteria during cultivation increases resistance to environmental stresses including resistance to low pH conditions [75,76], which might be at least partially mediated by cell wall modifications [75,77]. Regulation of pH may be used in industrial production to increase yields upon harvesting (i.e., enhance biomass yield per liter). Base addition prevents bifidobacterial end-products (e.g., acetic acid and lactic acid) to inhibit further growth of the organism, and thereby support increased cell density at harvesting [83]. To obtain high yields at the end of the fermentation, while maintaining resistance to acid stress, an intermediate pH control strategy was investigated (called partial pH control hereafter), where base addition is maintained until the end of the exponential phase and subsequently stopped to allow pH to drop rapidly as the cells enter the stationary phase.

It is important to understand how different growth conditions may influence the molecular make-up of probiotics strains [84]. Therefore, we assessed whether different pH regulation strategies affect the physiology of the *B. longum* NCC 2705 strain. For that purpose, we grew NCC 2705 under no pH control (no regulation), partial pH control (pH control until the end of the exponential phase) and full pH control (continuous pH control throughout the fermentation) and studied the impact of those different pH regulation regimes on the production of serpin. In addition, we investigated if other physiological traits (membrane permeability and acid resistance) were affected. Our study demonstrates that in NCC 2705, different growth conditions aiming at maximizing the final biomass yield influence the strain physiology, highlighting the need to better understand the impact of the production procedures traditionally geared to optimize final CFU counts.

Results

pH regulation during growth not only increases biomass yield but modifies serpin expression in *B. longum* NCC 2705. Here we examined how different pH regulation regimes affect the production of serpin in *B. longum* NCC 2705. For that purpose, we cultured *B. longum* NCC 2705 in a semi-defined medium containing glucose as sole carbon source (3%) and applied different pH regulation regimes (no pH regulation, partial pH control, full pH control). Samples were collected in early stationary phase to evaluate both yield and serpin levels. Partial or full pH control significantly increased the CFU levels approximately 10-fold as compared to the non-pH-controlled condition (2.4E8 CFU/ml), reaching final counts per ml of 2.5E9 and 2.7E9, respectively (Figure 1). However, low serpin levels were found in the cultures obtained under no and partial pH control (665 and 287 pg/mg total protein, respectively) while they were significantly higher when the pH was maintained at 6 throughout the entire fermentation (3144 pg/mg total protein). Overall, these results confirmed that pH regulation supports a higher biomass yield and demonstrated that, depending on the pH control regimen, they also modify the cell associated serpin levels produced by *B. longum* NCC 2705.

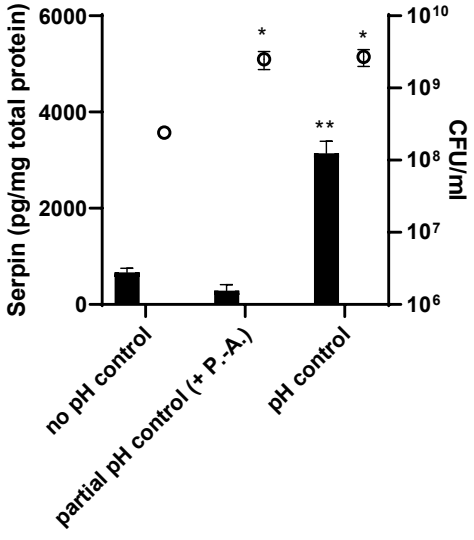


Figure 1: Yield (Colony Forming Unit per ml, circles) and serine protease inhibitor levels (black bars) reached in stationary phase *B. longum* NCC 2705 cells grown under different pH regulation (no, partial, full) regimes. Values represent average and standard deviations of biological and technical duplicates.

Full pH control drives glucose depletion, thereby activating serpin production. We have previously shown that glucose inhibits the production of serpin in *B. longum* NCC 2705 [140]. It is hence particularly surprising that in a fermentation using 3% of glucose as sole carbon source, full pH control enabled the production of serpin. Therefore, we determined whether the different pH regulation regimes led to distinct glucose consumption profiles that further impact serpin production. To this end, samples of *B. longum* NCC 2705 grown in the semi-defined glucose based medium under either partial or full pH regulation regime were collected at regular intervals (i.e., every hour). Optical density at 600 nm, pH, glucose levels in the culture supernatants and cell associated serpin levels were determined throughout the fermentations. In the partially pH controlled condition, maintaining pH at 6 by base addition during the exponential phase allowed to reach a high cell density (OD 10 after 8h), after which base addition was stopped to allow a fast acidification, i.e., the pH dropped from 6 to 4.58 in 3-4 hours (Figure 2 A). Under these conditions, a large amount of glucose was consumed (72 mM) during the exponential phase (i.e., within the first 8h of growth), but depletion only occurred later when the culture reached the stationary phase. Minimal levels of serpin were observed throughout the fermentation (about 207-277 pg/mg total protein) (Figure 2 A). Controlling the pH at 6 during the whole fermentation enabled to reach an even higher cell density (maximal OD of 17) and, in contrast with the partially controlled pH condition, glucose depletion happened already before the end of the exponential phase that appeared extended in this condition (i.e., +2 hours as compared to partial-pH control). From the point of glucose depletion (10h) onwards, serpin protein levels increased, finally reaching levels above 4000 pg/mg total protein during the early stationary phase (Figure 2 B). Of note, upon glucose depletion in the pH-controlled reactors, cell density decreased slightly (from OD_{600} 17 to 14), which might reflect a cell size reduction or lysis from that point in time (Figure 2 B). As we demonstrated previously, serpin production in *B. longum* NCC 2705 is only initiated in the absence of glucose, but is stimulated by the presence of an inducing sugar like galactose or fructose [140]. The semi-synthetic medium we used contains 3 % of yeast extract. The latter may vary in its content in residual sugars as for example Bacto™ and Difco™ low dusting yeast extracts contain 16 and 0.8 % of carbohydrates, respectively. The hypothesis that these residual sugars may interfere with serpin induction was corroborated by the fact that using a yeast extract with low residual carbohydrate content (i.e., Difco™ low dusting) led to substantially lower levels of

serpin (1645 vs 6800 pg/mg total protein, respectively) (Figure S1). Taken together, these results demonstrated that maintaining a pH of 6 throughout the entire fermentation enabled the depletion of glucose by metabolically active cells (i.e., during exponential phase), which depending on the yeast extract used allowed these cells to achieve significant serpin production.

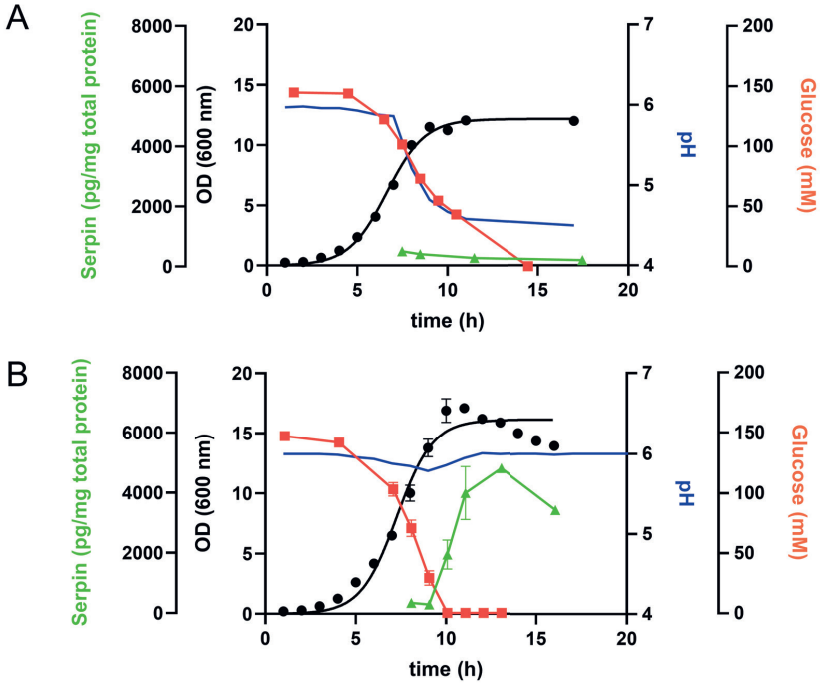


Figure 2: Fermentation outcomes of *B. longum* NCC 2705 grown in glucose (3%) under partial pH controlled (A) or maintained constantly at a pH of 6 (B). Growth is depicted using optical density measured at 600 nm (black dots) and its corresponding logistic growth model fitted curve (black line). Measured pH through the experiment (blue line), as well as glucose levels in the culture supernatant (red squares) and serpin levels (green triangles) are shown.

Variation in pH regulation regime impacts *B. longum* NCC 2705 cell membrane permeability. To understand the potential impact of different pH regulation regimes on cellular integrity and membrane permeability of *B. longum* NCC 2705, frozen mid-exponential and early stationary cells were subjected to propidium iodide (PI) staining and flow cytometry assessment. PI penetrates bacterial cells that have a compromised membrane integrity (e.g., are leaky), and further binds to their intracellular DNA. We confirmed that the conservation procedure (i.e., freezing in a glycerol containing buffer) did not impact cell integrity, as the permeability of both fresh and frozen cells was the same (data not shown). Furthermore, no difference in permeability was observed between cells harvested at mid-exponential or stationary phase when cultured under a partial pH-controlled regime, where in both growth stages cells displayed a relatively high mean fluorescence level (19050 or 18347, respectively) (Figure 3 A). Similar mean fluorescence level (18297) was observed in mid-exponentially harvested cells grown under full pH control. However, the fluorescence signal emitted by cells harvested from the stationary phase under full pH control conditions (pH constant at 6) was drastically lower (mean value 2324; Figure 3 B), indicative for a much higher membrane integrity (lower membrane leakiness) in these cells compared to the other three cell preparations.

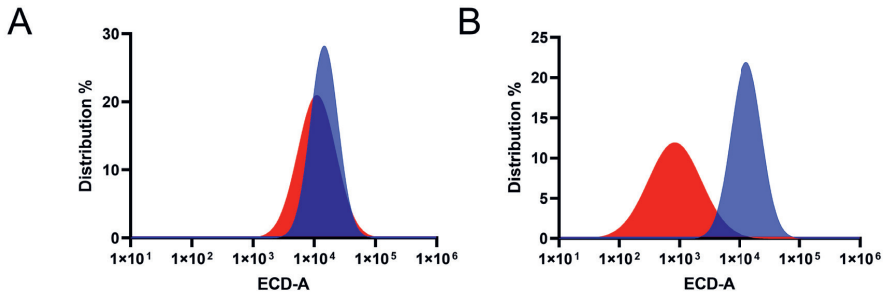


Figure 3: Membrane permeability measurements by flow cytometric analysis of PI stained *B. longum* NCC 2705 cultures grown under partial (A) or full (B) pH control. Population distribution (log2) of cells harvested during mid exponential (blue) or stationary (red) phases of growth are depicted.

Full pH control improves the strain resistance to low pH. An improved cell integrity (i.e., decrease in membrane permeability) in bacteria is associated with an enhanced resistance to low pH [192]. Therefore, we aimed at assessing whether permeability differences correlated with differences in resistance to low pH. To this end, we applied a double staining using 5-carboxyfluorescein diacetate acetocymethyl ester (CFDA) and PI [193] to assess using flow cytometry both the metabolic activity (esterase) and permeability of NCC 2705 exposed to an acidic pH. Initial results revealed the anticipated lower acid resistance of cells harvested during the exponential phase of growth as compared to stationary collected cells upon exposure to pH 2.9 for 18 minutes (Figure S2). Additionally, stationary cells grown under full pH control initially contained higher levels of metabolically active (CFDA+) cells (72%) relative to stationary cells grown using a partial pH regulation (20%). Moreover, the fraction of damaged cells (CFDA+/PI+) in the partially controlled pH culture was substantially larger than in the culture obtained using a full pH control (65% vs 15%, respectively) (Figure 4 A and B). This difference between the two conditions was maintained throughout the entire low-pH exposure period. At the end of the low pH incubation period (i.e., after 18 min), a remaining subpopulation of metabolically active cells (16%) was still present in the culture sampled from the fully pH regulated bioreactor. After the same incubation period, the metabolically active cell fraction derived from the partial pH-controlled culture almost completely disappeared (1%), being substituted by damaged (43%) and dead (55%) cells populations (Figure 4 A and B). Notably, the rate of increase of the dead cell fraction was comparable between the two conditions (respectively k=0.07766 and 0.08411) (Figure 4 C), hence supporting the initial metabolic and permeability differences to be the main drivers of the observed acid-stress resistance between the two cultures. These results indicate that NCC 2705 cells that were cultured using a constant pH of 6 could withstand low pH for a longer time as compared to cells obtained from a partial pH control culturing regime, therefore confirming the association between permeability and acid resistance.

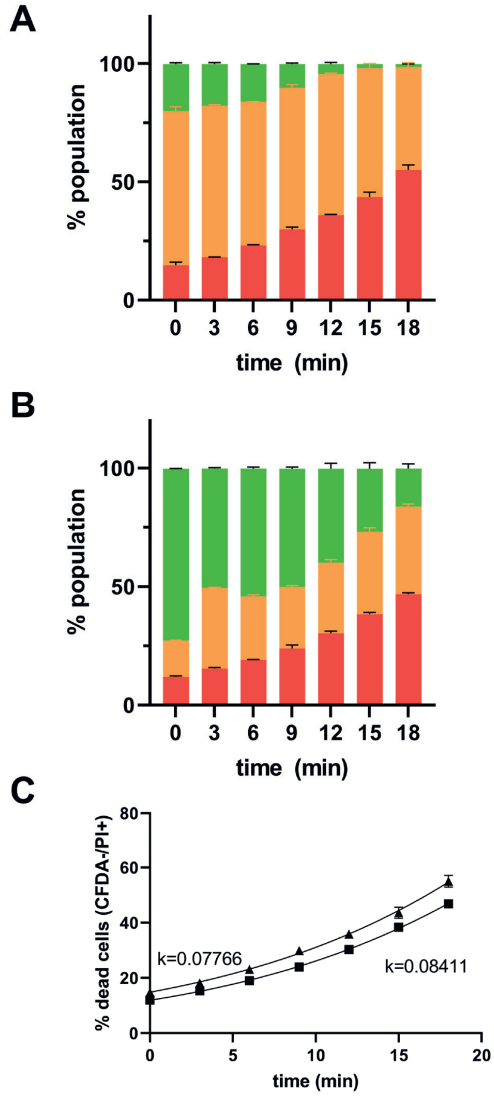


Figure 4: Evolution of early-stationary cultures of *B. longum* NCC 2705 upon exposure to a pH of 2.9 for 18 minutes. The distribution of live (CFDA+/PI-; green bars), damaged (CFDA+/PI+; orange bars) and dead (CFDA+/PI-; red bars) cell populations issued from partially pH-controlled (panel A), or full pH-controlled (panel B) cultures is depicted. Panel C represents the evolution of the dead cell populations through time in both cultures (partial pH control = triangles; full pH control = squares). Logistic growth model was applied, and statistical analysis was performed using a one-way ANOVA followed by a multiple test on the obtained increase rate (k).

Conclusions & discussion

We confirmed that in a semi-defined medium, pH control increases substantially the biomass yield at harvesting. In addition, our results demonstrate that different pH regulation regimes have various effects on the expression of serpin in *B. longum* NCC 2705. Base addition to control pH throughout the entire fermentation process resulted in a significant increase of the serpin protein exposed at the bacterial surface, which was not observed when base addition was stopped towards the end of the exponential phase.

We established that the pH mediated serpin induction was actually controlled by depletion of glucose from the medium, corroborating our earlier findings that glucose represses serpin production [140]. Maintaining a pH regulation through the entire fermentation slightly extended the exponential phase and led to a complete consumption of glucose before the end of this period, which subsequently led to induction of serpin production. Conversely, stopping the base addition towards the end of the exponential phase delayed glucose depletion which was only reached several hours after the entry in the stationary phase where the pH had already dropped considerably, i.e., to pH 4.5. Lactate and acetate are the main bifidobacterial fermentation by-products [21]. At a temperature of 25°C, acetate dissociates at a pH of 4.76 [194] meaning that below this pH acetate becomes protonated to form acetic acid. Under low pH (such as pH 4.5), the protonated acid may hence enter the cells by passive diffusion, where it will dissociate and increase the intracellular acidity leading to dissipation of membrane potential and loss of membrane integrity, acceleration of metabolic disfunction, and stagnation of protein biosynthesis [195]. This could explain the much lower capacity of the cells to produce serpin at that stage as well as the cell integrity loss observed under these conditions. Complementing our previous results [140] , this study showed that depending on the pH regime applied during cell culture, production of high levels of serpin can only be achieved if glucose depletion occurs while the cells are still metabolically active.

Our results underpin that presence of glucose strictly represses serpin production in metabolically active cells. Previous data demonstrated nevertheless that the presence of galactose and fructose facilitated an enhanced production of this probiotic effector molecule [140]. This suggests that another carbohydrate, potentially consumed or

imported upon glucose depletion, is likely required to reach the high level of serpin we observed in the full pH regulated bioreactors. To fulfill the bifidobacterial nutritional requirements, without having to develop complex defined media [78], yeast autolysates (called yeast extracts) and hydrolyzed proteins (called hereafter peptones) are often added as a source of nitrogen, vitamins and other micronutrients [79]. Yeast extracts and peptones are complex ingredients, and their source or manufacturing process influences the level of amino acids, vitamins, minerals, and carbohydrate they contain [80-82]. In this work, we observed that yeast extracts with different carbohydrate contents (high or low) led to different levels of serpin induction under full pH regulation (Figure S 1), supporting the hypothesis that in addition to glucose depletion, the carbohydrates contained in these complex growth medium ingredients may “amplify” serpin production in *B. longum* NCC 2705 once glucose is depleted.

Cell staining coupled to flow cytometry indicated that different pH regulation regimes affect the permeability of the bacterial cells. As expected, a relatively low cell integrity was observed during the logarithmic phase of growth, which improved solely during the stationary phase in *B. longum* NCC 2705 cultured at a constant pH of 6. This improvement in cell integrity did not occur in the partial pH control regime where the pH dropped abruptly at the end of the exponential phase. The improved cell integrity correlated with an improved resistance to low pH, which agrees with previous observations [192]. It was previously demonstrated that acid-stress exposure induces stress adaptation responses that mediated protection against lethal acid stress challenge conditions [75,196]. In this context our data are somewhat counterintuitive as it is surprising that cells that were cultured at a constant pH of 6 (i.e., have not been exposed to acidic conditions) displayed an increase pH tolerance relative to cells derived from an acidified culture.

In our experiments, we stopped base addition towards the end of the exponential phase, which resulted in relatively fast acidification with a pH drop to 4.5 within 3-4 hours. As mentioned above, at this pH acetate becomes protonated [194] and likely causes major disruptions in the cell homeostasis and biosynthetic capacities [195], which we propose impaired the induction of the stress response observed in previous studies [75,196]. On the other hand, the glucose depletion observed under full pH regulation happened while the cells were metabolically active, at a pH were all produced fermentation by-products stay non protonated. This may have promoted the cells to

transition to another carbohydrate present in the yeast extract or peptone, which shifted the bacteria to a different physiological state. Overall, this supports the idea that different carbohydrates may modulate the physiology of *B. longum* NCC 2705, rendering the bacteria less permeable and more acid resistant. The exact mechanism underlying this effect remains to be unraveled.

In conclusion, we showed that different pH regulation regimes aiming at increasing bacterial yields may impact the physiology of *B. longum* NCC 2705. We suggest that the most striking changes in the physiology of the strain grown under a constant pH of 6 (serpin levels, membrane permeability and increased pH resistance) are driven by changes in the metabolized carbohydrate during fermentation. Besides the carbohydrate mediated regulation of serpin [140], several studies have highlighted other bifidobacterial effector molecules such as the *pil2* and *pil3* sortase dependent pili of *B. bifidum* PRL2010 [131] to be subjected to carbohydrate mediated regulation [129,135]. Taken together, our results underline the importance of carbohydrate consumption patterns in bifidobacterial fermentations and illustrate the need to improve our understanding of their potential influence on the functional molecular make-up of the bacterial cells produced.

Materials and methods

Strain & culture conditions. *Bifidobacterium longum* NCC 2705 was retrieved from the Nestle Culture Collection (NCC) (Nestlé Research, Lausanne, Switzerland). Growth was routinely performed using de Man, de Rogosa, Sharpe medium (BD Difco, Franklin Lakes, USA) supplemented with 0.05 % of cysteine (MRSc), at 37°C in anaerobiosis without agitation. Growth under controlled conditions was performed in 1 L DasGip bioreactors (Eppendorf, Hamburg, Germany). In this case, preculture and cultures of NCC 2705 were grown in a semi-defined medium (SM), consisting of 0.15% Tween-80, 3 % glucose (Sigma-Aldrich Chemie GmbH, St-Louis, USA), 3 % Bacto™ Yeast Extract and 2% Bacto™ peptone (Thermo Fisher Scientific AG, Basel, Switzerland). An alternative low carbohydrate containing yeast extract, the Difco™ low dusting (Thermo Fisher Scientific AG) was used at 3% instead of the Bacto™ Yeast Extract (see supplementary data). According to the supplier data, the total carbohydrate content of the Bacto™ Yeast extract is of 163.3 mg/g, while in both the Difco™ low dusting yeast extract and Bacto™ peptone, this amount is negligeable (8.2 and 6.3 mg/g, respectively). Inoculation of the main cultures was performed at an equivalent Optical Density (OD) at 600 nm of 0.1. Temperature was set at 37°C, agitation at 200 rpm and gazing was performed with CO_2 in the headspace. Acidification in the different bioreactors was followed using online connected pH probes. No base was added for the non-pH-controlled condition, while 30% NaOH solution (Sigma-Aldrich Chemie GmbH) was supplemented to maintain a pH of 6, until end of exponential phase (referred to as partial pH control) or until the end of the fermentation (referred to as full pH control). At regular intervals throughout the entire growth experiment, samples were collected and used to measure OD at 600nm, before being centrifuged (3500 g, 2 min, 4°C), resuspended in PBS supplemented with 15% sterile glycerol and stored at -80°C. Frozen samples were then further used for Colony Forming Units (CFU) determination on MRSc agar while glucose concentration was measured in the culture supernatants. Every fermentation condition was run in biological duplicates.

Glucose measurement by HPLC. Glucose level was quantified in filtered sterile culture supernatants by High Performance Liquid Chromatography (HPLC, Agilent Technologies AG, Basel, Switzerland), using an Aminex HPX-87H columm (Bio-Rad Laboratories AG, Cressier, Switzerland). 5 µl of samples were injected, using a 5 mM H_2SO_4 solution as eluent with a flow rate of 0.6 ml/min (run time 25 min). Chromatogram was obtained using a refractive index detector, from which peak areas and corresponding concentrations were calculated.

Preparation of crude extracts & total protein quantification. Frozen samples were collected and centrifuged (3500 g, 2 min, 4°C) to obtain bacterial pellets equivalent to 1 ml of culture. Obtained pellets were washed once with one volume of ice-cold PBS and resuspended in 960 µl of PBS containing SigmaFast Protease Inhibitor (Sigma-Aldrich Chemie GmbH; 1 tablet per 100 ml of solution). The obtained solution was subsequently lysed by bead-beating using a FastPrep-24 (MP Biomedicals SARL, Illkirch-Graffenstaden, France), 3 times for 1 minute at 4 m/s, with 2 minutes cooling on ice in between. Crude lysates were used as such, containing both soluble and non-soluble fractions. Total protein content was determined using the Pierce BCA protein Assay kit (Thermo Fisher Scientific AG).

Serpin protein quantification by sandwich ELISA. Anti-serpin sandwich ELISA was performed as described by Duboux et al [140]. Briefly, 96-well plates (Nunc MaxiSorp [Thermo Fisher Scientific AG]) were coated with anti-serpin antibodies. After washing and blocking of the wells, samples were applied and a titration range of purified recombinant serpin [61] was used to generate a standard curve. Bound serpin protein was detected using biotinylated anti-serpin antibodies. Detection was performed by addition of streptavidin-labelled horse-radish peroxidase (HRP Pierce, Thermo Fisher Scientific AG) enzyme conjugate, followed by washing to remove background reactivity. Chromogenic detection was performed by addition of HRP substrate (1-step Ultra TMB ELISA, ThermoFisher Scientific AG) followed by 15 minutes incubation. The reaction was terminated by addition of 100 µl of 2M sulfuric acid followed by optical density at 450 nm measurement in a Varioskan spectrophotometer (Thermo Fisher Scientific AG).

Flow cytometry & stainings. Flow cytometry analyses were performed with a Beckman Coulter Cytoflex S (Beckman Coulter, Brea, US) equipped with four lasers. Bacterial cell staining was performed using 1.5 µM propidium iodide (PI, Thermo Fisher Scientific) and 5 µM carboxyfluorescein diacetate acetocymethyl ester (CFDA, Thermo Fisher Scientific) with an incubation of 10 minutes at room temperature before analysis. Acquisition threshold was set at 1000 in order to capture cells of all sizes. Cell size was estimated using the Forward Scatter signal (FSC). Cell permeability was evaluated through PI staining and measured using the ECD channel (excitation at 561nm, emission at 610/20 nm). Metabolic activity was evaluated using the cell-permeant esterase substrate CFDA measured using the FITC detector (excitation at 488 nm, emission at 525/40 nm). Acquisitions were processed with the Beckman-Coulter CytExpert Acquisition software version 2.3, followed by further analysis with the FCSExpress software version 7.01 (De Novo Software, Pasadena, USA). Statistical analysis of population distribution was performed in GraphPad Prism (v8.1.1, GraphPad Software Inc., San Diego, USA), using a welch-t test.

Metabolic activity loss upon pH 2.9 exposure. Frozen mid-exponentially or stationary grown cultures of *B. longum* NCC 2705 conserved in PBS 15% glycerol at -80°C were used as starting material. Cells were thawed on ice, washed in one volume of sterile PBS (centrifugation 2 min, 3300 g), and resuspended to an OD of 1.0 in PBS adjusted at a pH of 2.9 using 1 M hydrogen chloride (HCl). The obtained solution was distributed in 100 µl aliquots and incubated at 37°C for a period of 18 minutes. An aliquot was collected every 3 min, and 900 µl of PBS containing 1.5 µM PI and 5 µM CFDA was added to stop the reaction. After an incubation of 10 minutes at room temperature, permeability, and metabolic activity of 5000 individual cells was measured by flow cytometry (see above) at each timepoint and for each condition. Fractions of live (CFDA+/PI-), damaged (CFDA+/PI+) and dead (CFDA-/PI+) cell contained within each culture at each time point were retrieved. Dead cell populations increase over time was modelled in GraphPad Prism (v8.1.1, GraphPad Software Inc) using a logistic growth curve, as recommended by de Besten et al. [197]. Assay was performed in biological duplicates and include technical duplicates in each test. Additionally, frozen cells were compared to freshly obtained cultures to ensure that the freezing process did not impact the above metabolic activity readouts.

Supplementary figures

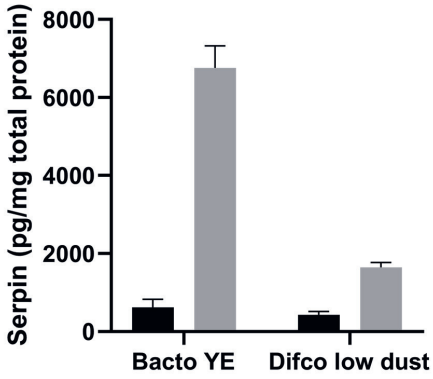


Figure S 1: Serpin levels produced by *B. longum* NCC 2705 grown using a high (Bacto YE) or low (Difco low dusting) carbohydrate containing yeast extract. Levels reached in non pH regulated (black bars) or pH 6 controlled (grey bars) fermentations are depicted.

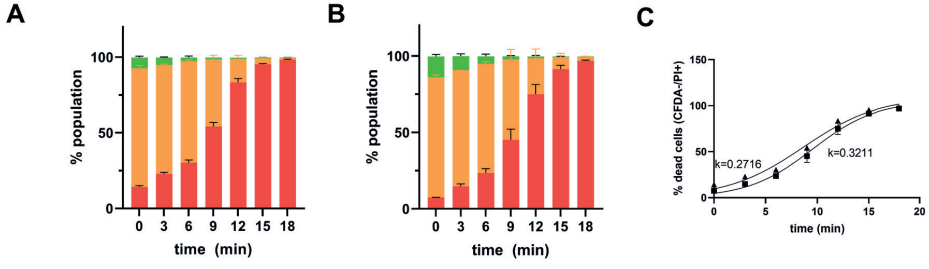


Figure S 2: Evolution of mid-exponential cultures of *B. longum* NCC 2705 upon exposure to a pH of 2.9 for 18 minutes. The distribution of live (CFDA+/PI-; green bars), damaged (CFDA+/PI+; orange bars) and dead (CFDA+/PI-; red bars) cell populations issued from partially pH-controlled (panel A) or full pH-controlled (panel B) cultures is depicted. Panel C represents the evolution of the dead cell populations through time (partial pH control = triangles; full pH control = squares). Logistic growth model was applied, and statistical analysis was performed using a one-way ANOVA followed by a multiple test on the obtained increase rate (k).

CHAPTER 5

5

Using fluorescent promoter-reporters to study sugar utilization control in *Bifidobacterium longum* NCC 2705

S. Duboux [1,2] | J.A. Muller [1] | Filippo De Franceschi [1] |
A. Mercenier [2] | M. Kleerebezem [2]

[1] Nestlé Research, Route du Jorat 57, CH 1000 Lausanne 26, Switzerland
[2] Host-Microbe Interactomics Group, Wageningen University & Research, De Elst 1, 6708WD Wageningen, The Netherlands

This chapter was published as Duboux, S., Muller, J. A., De Franceschi, F., Mercenier, A. & Kleerebezem, M. Using fluorescent promoter-reporters to study sugar utilization control in Bifidobacterium longum NCC 2705. Sci Rep 12, 10477, doi:10.1038/s41598-022-14638-4 (2022)

Abstract

Bifidobacteria are amongst the first bacteria to colonize the human gastro-intestinal system and have been proposed to play a crucial role in the development of the infant gut since their absence is correlated to the development of diseases later in life. Bifidobacteria have the capacity to metabolize a diverse range of (complex) carbohydrates, reflecting their adaptation to the lower gastro-intestinal tract. Detailed understanding of carbohydrate metabolism regulation in this genus is of prime importance and availability of additional genetic tools easing such studies would be beneficial. To develop a fluorescent protein-based reporter system that can be used in *B. longum* NCC 2705, we first selected the most promising fluorescent protein out of the seven we tested (i.e., mCherry). This reporter protein was then used to study the carbohydrate mediated activation of P_{Bl1518} and P_{Bl1694}, two promoters respectively predicted to be controlled by the transcriptional factors AraQ and AraU, previously suggested to regulate arabinose utilization and proposed to also act as global transcriptional regulators in bifidobacteria. We confirmed that in *B. longum* NCC 2705 the AraQ controlled promoter (P_{Bl1518}) is induced strongly by arabinose and established that the AraU controlled promoter (P_{Bl1694}) was mostly induced by the hexoses galactose and fructose. Combining the mCherry reporter system with flow cytometry, we established that NCC 2705 is able to co-metabolize arabinose and glucose while galactose was only consumed after glucose exhaustion, thus illustrating the complexity of different carbohydrate consumption patterns and their specific regulation in this strain.

Introduction

Bifidobacterium is an important genus within the human intestinal microbiota. Bifidobacterial species are among the first to establish in the neonate human gastrointestinal tract (GIT) and are predominant in the breast-fed newborn until the age of 4 months [7]. Their presence in early life has been proposed to play a critical role in the maturation of the newborn's immune system as several studies have linked low abundance of members of this genus in the infant gut to later in life immune disorders such as atopy [38] or asthma [39]. Early-life *Bifidobacterium* dominance has also been proposed to be linked with reduced risks of obesity [40] or acquisition of antimicrobial resistances [41] later in life.

The genome of bifidobacteria encodes an extensive enzyme repertoire that supports their capacity to effectively breakdown and metabolize diverse carbohydrate substrates [16]. The diversity of carbohydrates that can be metabolized by bifidobacteria is well reflected within the *B. longum* subspecies. Strains of *B. longum* subsp. *infantis* have the particular ability to hydrolyze several of the indigestible human milk oligosaccharides (HMOs) and to metabolize the derived galactose and fucose, explaining their abundance in the intestine of breast-fed infants [8]. Strains of *B. longum* subsp. *longum* can also metabolize carbohydrates found in human milk such as lactose [17] or Lacto-N-Tetraose (LNT), but they appear to be particularly well equipped to import and consume food fibers derived carbohydrates, as reflected by the genome of *B. longum* subsp. *longum* NCC 2705 strain [19]. Bifidobacteria can be stimulated by the consumption of arabinose-containing cereal fibers [21]. This type of high molecular weight fiber (i.e., arabinogalactans) is first degraded by specific members of the microbiota community such as *Bacteroides caccae,* producing smaller arabinose containing oligosaccharides that can further be consumed by secondary degraders as demonstrated for *B. longum* NCC 2705 [33]. In *Bifidobacterium*, all carbohydrates are theoretically processed through the relatively unique central carbon metabolism that includes the so-called "bifid-shunt", which involves the bifunctional xylulose 5-phosphate/fructose-6-phosphate phosphoketolase enzyme [21]. As an example, in *B. longum* NCC 2705, glucose, fructose, mannose, xylose, ribose and galactose were all shown to be catabolized via the "bifid-shunt" [44].

Since carbohydrate metabolism is probably their predominant source of ATP, adequate control of carbon import and metabolic flux is important in the physiology of *Bifidobacterium* species [51]. This is reflected by the fact that bifidobacterial genomes contain a relatively large number of genes that encode carbohydrate binding Lac-I type transcription regulators, which are often predicted to regulate genes located in their close vicinity [182]. As an example, the ROK-family regulator AraU and the LacI-family regulator AraQ were predicted to control adjacent operons encoding putative arabinose transporters and catabolic enzymes [198]. However, in *B. breve* UCC 2003 AraQ, together with another Lac-I type regulator MalR1, were demonstrated to control a large set of genes that are spread over the genome including not only genes involved in uptake and metabolism of various carbohydrates, but also genes involved in the ”bifid-shunt” as well as several transcription regulators [181]. We recently demonstrated that the production of serine protease inhibitor (serpin) of *B. longum* NCC 2705, a key bifidobacterial effector molecule, is tightly controlled by carbohydrates and established that the presence of glucose in the growth medium inhibits its production [140]. These results suggest that the canonical process known as glucose-catabolite repression is important in controlling carbohydrate utilization by bifidobacteria, although the underlying molecular mechanism and involved regulatory factors remain to be deciphered. Lac-I type regulatory proteins represent the closest homologues (24-32% identity) to the catabolite control protein A (CcpA) implicated in the glucose-catabolite repression in many Gram-positive bacteria genes [179]. Of note, *B. longum* NCC 2705 harbors 16 genes encoding Lac-I type regulators according to the RegPrecise database [180].

The poor genetic accessibility of bifidobacteria has hampered the unambiguous analysis of the mechanisms underlying carbohydrate regulation in members of this genus. Progress has been made over the past decades, especially in improving electroporation protocols [199] and elaborated solutions were developed to enable genetic modifications of recalcitrant *Bifidobacterium* strains [200]. Nevertheless, most available plasmid vectors have a relatively narrow replication host-range, with pMDY23 being a notable exception as it could replicate in nine different species and subspecies [200]. Altogether, validated gene-expression and promoter-reporter systems remain relatively scarce for this genus [200,201]. To date, *Bifidobacterium* promoter activity studies mainly used enzyme reporters such as the β-glucuronidase encoded

by the *gusA* gene [201-203]. Over the past decades, fluorescent protein reporters have been extensively employed to obtain high-throughput and single cell resolution expression information in both Gram-negative and Gram-positive bacteria [204-206], including several lactic acid bacteria [207,208]. The Cyan (CFP), green (GFP), yellow (YFP) and red (mCherry) fluorescent proteins that all require oxygen for their maturation as well as the anaerobic maturing cyan-blue Evoglow-Pp1 fluorescent protein have been functionally expressed in different bifidobacterial species [209-211], but were mainly used up to now to track the presence of labelled cells in different environments [210-212] rather than studying promoter regulation.

In the present study, we first describe the selection of an appropriate fluorescent reporter protein and its validation using *B. longum* NCC 2705 promoters of different strengths that were previously characterized using the β-glucoronidase (*gusA*) based reporter [203,212]. We then demonstrated that mCherry can be used to decipher the activity and activation patterns of two promoters (P_{Bl1518} and P_{Bl1694}) controlled by the AraQ and AraU transcriptional regulators, which were predicted to be implicated in arabinose utilization but also proposed to act as global transcriptional regulators in bifidobacteria. Taking advantage of the single cell resolution of fluorescent protein-based reporters we further established that when combined with flow cytometry, mCherry enables the measurement of (sub-)population adaptation during carbohydrate substrate transition and underpins the distinction between sequential utilization and co-consumption of different sugars present in the medium. Altogether, our data demonstrate the value of fluorescent protein reporters in shedding light on the regulation of carbohydrate metabolism in *B. longum* NCC 2705, providing a detailed resolution of the way this strain behaves in a complex substrate environment.

Results

Selection and validation of fluorescence-based reporters for promoter activity measurement in *B. longum* NCC 2705. We first aimed at selecting a reporter displaying an optimal fluorescence signal. For that purpose, the genes encoding fluorescent proteins that either require oxygen for their activity (GFP, CFP, YFP, mCherry) or are fluorescent under anaerobic conditions (Pp1, Bs1, Bs2) were cloned and expressed in *B. longum* NCC 2705 under control of the strong P_{gap} promoter that drives the expression of glyceraldehyde 3-phosphate dehydrogenase in *B. bifidum* S17. Recombinant *B. longum* NCC 2705 producing the GFP & YFP proteins emitted a fluorescent signal that was distinguishable from the autofluorescence detected in the wild-type cells, while the specific detection of CFP in the corresponding producing strain was hampered by strong autofluorescence displayed by NCC 2705 (Figure S 2). The strain harboring the pVG-mCherry (Table S 1) produced the strongest relative fluorescence signal, which in part was attributable to the lack of autofluorescence produced by the wild-type cells at the corresponding excitation and emission wavelengths (Figure S 2). Amongst fluorescent proteins that do not require oxygen exposure for their maturation, Pp1 expressed in *B. longum* NCC 2705 under control of the P_{gap} promoter (pSDU03) generated the highest intensity signal relative to the mild autofluorescence detected for the wild-type cells (Figure S 3).

Next, to validate that fluorescent proteins can be used to monitor different levels of promoter activity in *B. longum* NCC 2705, the genes encoding the best candidate fluorescent proteins (i.e., mCherry and Pp1) were cloned under control of the raffinose inducible P_{Bl1518} promoter [203,212]. This promoter and the P_{gap} promoter were previously characterized using the *gusA* reporter system and thereby provide a congruent set of promoter-reporter constructs for comparative analysis (Table S 1) [203,212]. Initial experiments using *B. longum* NCC 2705 derivatives that harbor the previously described *gusA* reporter constructs confirmed previously reported results [203,212], including the lack of activity of P_{Bl1518} in cells grown on glucose, the activation of this promoter during growth on raffinose (35-fold induced GusA expression), as well as the highest level of GusA activity detected in cells harboring the P_{gap}-*gusA* reporter construct (>5-fold higher activity compared to raffinose grown P_{Bl1518}-*gusA* harboring strain) (Figure 1 A). The fluorescence levels measured for *B. longum* NCC 2705 cultures harboring the mCherry or Pp1 genes under the control of the P_{gap} and P_{Bl1518}

promoters principally recapitulated the results obtained with the corresponding *gusA* reporter constructs (Figure 1 B & C). P_{gap} was confirmed to be the most active promoter generating the highest fluorescence levels with either the mCherry or Pp1 reporter constructs (Figure 1 B & C). Moreover, the induction of P_{Bl1518} during growth on raffinose compared to growth on glucose could be confirmed for both fluorescent protein reporters, albeit with a substantially lower fold-induction (3.5- and 3-fold for the mCherry and Pp1 constructs, respectively) compared to the β-glucoronidase reporter (Figure 1). This difference appeared to be attributable to background fluorescence levels detected in glucose grown cultures harboring the mCherry or Pp1 reporters cloned under control of the P_{Bl1518}, whereas the β-glucoronidase expression was barely detectable under these conditions. In addition, the wildtype strain displayed a relatively high autofluorescence at the Pp1 excitation and emission wavelengths, which may compromise accurate activity measurements of promoters with low levels of activity. Notably, similar high autofluorescence signals at the wavelengths used for Pp1 activity measurement were also detected in *B. longum* subsp. *infantis*, *B. breve*, and *B. animalis* subsp. *lactis*, which may limit the use of this fluorescent protein reporter in *Bifidobacterium* spp. (Figure S3). Importantly, barely detectable background autofluorescence levels were observed for the mCherry excitation and emission wavelengths in *B. longum* NCC 2705 (Figure 1) or the other *Bifidobacterium* spp. (Figure S3). Background autofluorescence levels seriously impact the assessment of promoter activity, which is obvious when comparing the P_{gap} activity assessed by the mCherry and Pp1 reporters relative to the autofluorescence background in *B. longum* NCC 2705 (600- and 16-fold, respectively; Figure 1) and *B. breve* ATCC 15700 (54- and 5-fold, respectively; Figure S 4).

Taken together, these results demonstrate that both mCherry and Pp1 can be used as promoter-reporter proteins in *B. longum* NCC 2705 and other *Bifidobacterium* spp.. Despite the fact that mCherry requires oxygen for its maturation, which limits its use to post-cultivation activity determinations, our results indicate that it provides a superior sensitivity and dynamic range as compared to Pp1. For these reasons, we decided to use mCherry as the fluorescent protein reporter of choice in our subsequent experiments (see below).

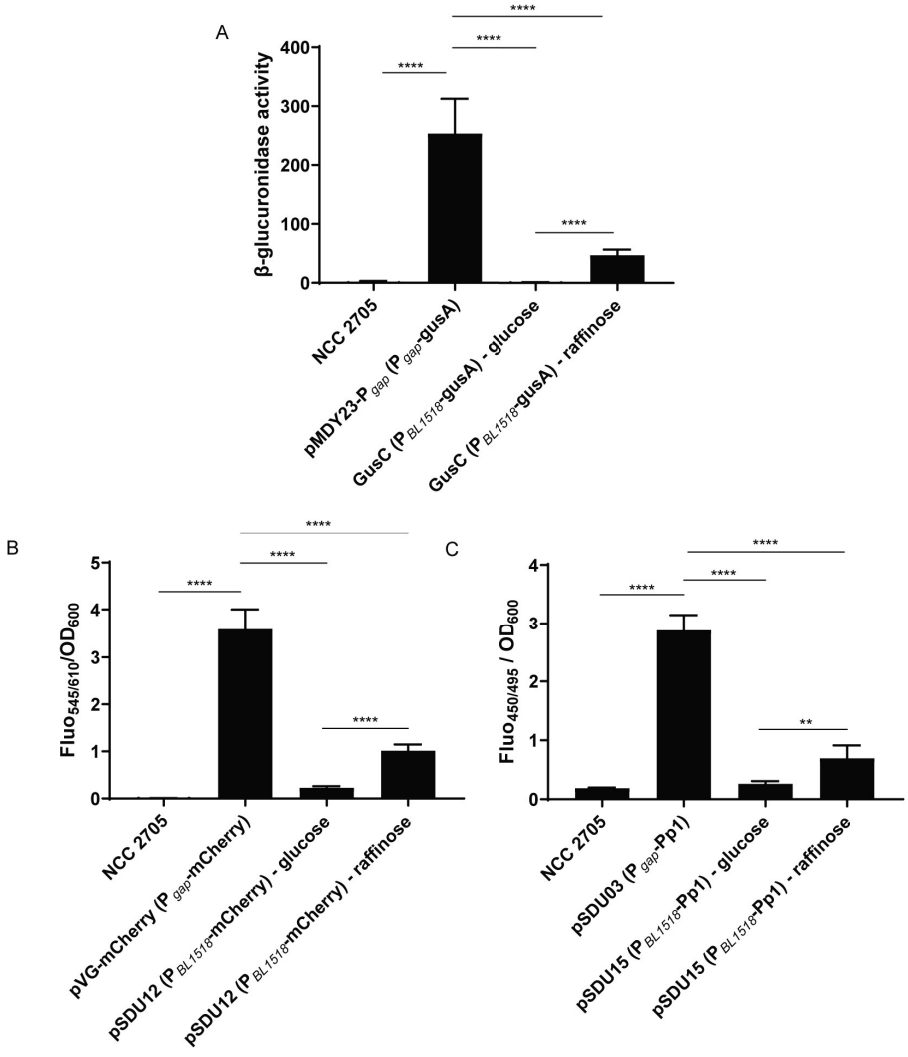


Figure 1: *B. longum* NCC 2705 promoter activities measured using different reporter systems in mid-exponentially growing cells. Background activity/fluorescence of the wildtype and recombinant strains harboring reporter genes under the control of P_{gap} (grown on glucose) or P_{BL1518} (grown on glucose or raffinose) was measured through β-glucoronidase (A), fluorescence emitted by mCherry (B) and fluorescence emitted by the anaerobic cyan-blue Pp1 (C). Values represent average of biological triplicates. Significant differences were calculated using one-way ANOVA, followed by a Sidak's multiple comparison test (** p-value <0.01; **** p-value < 0.0001).

Carbohydrate regulation of AraQ and AraU targeted promoters, P_{BL1359} and P_{BL1694}. The mCherry-based reporter system was used to study carbohydrate mediated regulation of promoter activity in *B. longum* NCC 2705 using two exemplary promoters. The first promoter sequence we selected is located upstream of the aldose-1-epimerase encoding gene (*galM; BL1359*) (Figure S 5 A). This promoter has been predicted to be regulated by AraQ (*BL0275*) [180], an arabinose binding Lac-I type regulator protein that is conserved in *Bifidobacterium* species [198]. AraQ is not only implicated in the regulation of arabinose metabolism but has also been reported to control genome-wide regulation of gene expression in bifidobacteria [181,182], including the closely located glyceraldehyde-3-phosphate dehydrogenase (*gap*) gene (Figure S 5 A). Glucose grown cells harboring the P_{BL1359}-mCherry reporter construct (pSDU30) did produce a 100-fold higher fluorescence level compared to the autofluorescence levels detected in untransformed wildtype cells (data not shown). The P_{BL1359} promoter was shown to be induced to its highest level when the strain was cultured on arabinose (3.6 fold induction compared to glucose grown cells) (Figure 2 A). It was also activated in cells growing on fructose, galactose and ribose (1.8, 1.75 and 2.2 fold induction, respectively), while no significatively higher promoter activity was observed in cells growing on xylose (Figure 2 A).

The second promoter region that was selected (P_{BL1694}) is located upstream of the *BL1694-BL1696* operon, which is annotated to encode a multiple sugar ABC transporter [19] (Figure S 5 B). Homologues of this transporter, annotated as AraFHG, were shown to be present in several Actinobacteria families and were proposed to be involved in arabinose import [198]. AraFGH encoding genes are widely distributed in Bifidobacteriaceae genomes and are predicted to be regulated by the adjacent but divergently transcribed *araU* gene (*BL1693*) that encodes a ROK-family transcriptional regulator [180,198]. It is worth to note that the region we cloned is predicted to contains two divergent AraU binding sites, controlling the expression of BL1693-1692 (P_{BL1693}) and BL1693-1695 (P_{BL1694}), respectively (Figure S 5 B) and that our cloning strategy specifically aimed at deciphering the activity of P_{BL1694}. While glucose grown cells harboring the P_{BL1694}-mCherry construct (pSDU32) exhibited a fluorescence level similar to that observed in glucose-grown cells harboring pSDU30 (P_{BL1359}-mCherry), P_{BL1694} activity was strongly induced when cells were grown on fructose and galactose (20- and 28-fold relative to the glucose expression level, respectively). Notably, the

BL1694 promoter activity was also stimulated during growth on the pentoses ribose and xylose, and during growth on arabinose albeit to a lower level compared to galactose or fructose (Figure 2 B).

These results establish that both P_{BL1359} and P_{BL1694} are controlled by carbohydrates, as they have a relatively low activity on glucose and are activated by growth on more than a single carbohydrate.

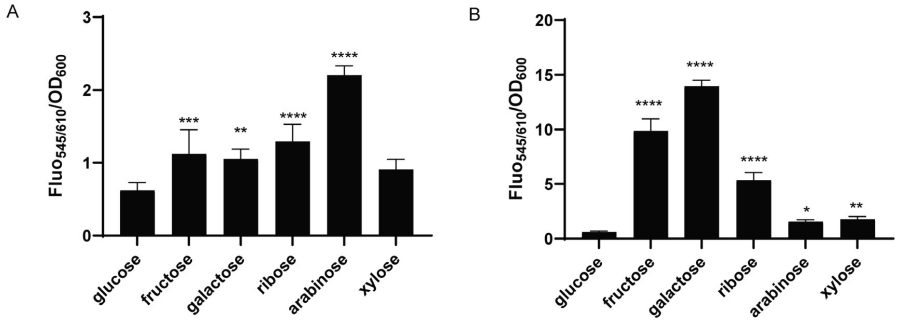


Figure 2: *B. longum* NCC 2705 promoter activities of mid-exponentially grown cells on 1% different hexoses and pentoses. Promoter activity is measured as the fluorescence emitted by the mCherry protein, normalized by the cell density (OD 600 nm). Fluorescence emitted by the cells harboring the pSDU30 (P_{BL1359}-mCherry) or the pSDU32 (P_{BL1694}-mCherry) construct is depicted in panels A and B, respectively. Values represent average of biological duplicates and technical triplicates. Significant differences were calculated using one-way ANOVA, followed by a Sidak's multiple comparison test (* p-value <0.05; ** p-value <0.01; *** p-value <0.001; **** p-value < 0.0001).

Population behavior of *B. longum* NCC 2705 switching from glucose to arabinose or galactose. We further investigated the role of glucose in the regulation of P_{BL1359} and P_{BL1694}. Moreover, we exploited the capacity of flow cytometry to detect single cell-fluorescence levels to assess community behavior during transition from one carbon source to another, e.g., from glucose to either arabinose or galactose.

We first confirmed that the mCherry reporter system we developed allows the detection and quantification of bacterial sub-populations. To this end, we performed a validation experiment with mixtures at different ratios of stationary phase cultures of *B. longum* NCC 2705 harboring pSDU30 (P_{BL1359}-mCherry) and pSDU32 (P_{BL1694}-mCherry) grown on glucose or arabinose, and on glucose or galactose, respectively. The different mixtures were analysed by flow cytometry, revealing an almost perfect separation and quantification of the high (arabinose or galactose grown, respectively) and low (the glucose grown counterparts) fluorescent populations present in the mixtures (Figure 3). This result demonstrates that the mCherry reporter system allows to distinguish sub-populations of cells growing on different sugars with very high accuracy, but also confirmed the bi-modal activation of P_{BL1359} and P_{BL1694}, both promoters being mostly repressed on glucose and activated on arabinose or galactose, respectively.

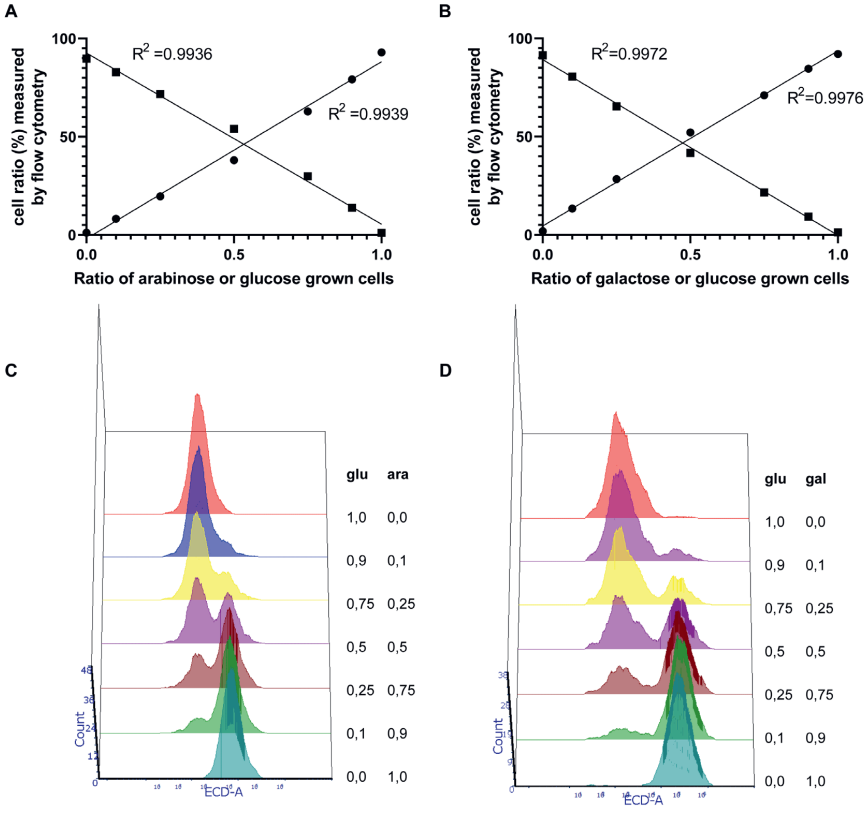


Figure 3: Flow cytometry scatter plots of stationary grown *B. longum* NCC 2705 cells harboring the pSDU30 (P_{BL1359}-mCherry) (panel A & C) and the pSDU32 (P_{BL1694} – mCherry) (panel B & D) constructs. Populations were mixed at different ratios (A & C [glucose:arabinose cells], B & D [glucose:galactose cells]) of 1.00:0.00, 0.90:0.10, 0.75:0.25, 0.50:0.50, 0.25:0.75, 0.10:0.90 and 0.00:1.00. Quantification of high (circles) and low (squares) fluorescent populations in the different mixes was performed by flow cytometry. Lines represent linear regression for each population (Panels A & B).

We next analysed sugar utilization and population adaptation of the recombinant *B. longum* NCC 2705 strains containing pSDU30 (P_{BL1359}-mCherry) or pSDU32 (P_{BL1694}-mCherry) during the transition from glucose to arabinose or from glucose to galactose, respectively. To this end, the NCC 2705 strains harboring pSDU30 or pSDU32 were grown on MRSc-C media supplemented with 0.1 % glucose and 0.5 % arabinose, or 0.1 % glucose and 0.5 % galactose, respectively. During growth, the consumption of the available sugars was followed while in parallel, single-cell fluorescence levels were measured in the population by flow cytometry. The results clearly demonstrated that pSDU30 harboring *B. longum* NCC 2705 consumed glucose and arabinose simultaneously during the initial hours of growth, which was supported by relatively high average levels of fluorescence in the culture that reflect substantial activity of the P_{BL1359} promoter as a proxy for expression of the arabinose utilization genes (Figure 4 A & B). After approximately 4 hours, glucose was depleted from the medium, which coincided with a modest but significant increase in average fluorescence in the culture, indicating that upon glucose depletion the activity of the P_{BL1359} promoter is further enhanced. This is in agreement with the observed acceleration (4.6X) of the rate of arabinose consumption after glucose was depleted (Figure 4 A), indicating that despite glucose being exhausted, cells are alive and metabolically active. Analogously, the fluorescence levels measured for single cells in the culture increased towards the level observed in cells grown on arabinose alone only after glucose was depleted (Figure 4 B & C). These results show that the presence of arabinose in the medium alleviates the observed glucose-suppression of P_{BL1359} activity, but that the activity of this promoter can still be further enhanced when glucose is depleted, indicating that the presence of glucose still mediates a partial level of repression in media that also contain the arabinose.

These results strikingly contrasted with those obtained for the strain harboring pSDU32, containing the P_{BL1694}-mCherry reporter. During initial growth the cells appeared to exclusively utilize glucose, which is in agreement with low average P_{BL1694} activity in the culture. Following the depletion of glucose from the medium (after approximately 3 hours), the strain remained alive and started to actively utilize galactose for growth which coincided with activation of P_{BL1694} promoter, reflected by the sharp increase of the average fluorescence levels in the culture (Figure 4 D). This typical diauxic growth characteristic on is in agreement with the single cell analysis,

which showed that initial fluorescence levels were similar to those observed in cultures grown on glucose alone and once glucose was depleted single cell fluorescence levels sharply increased to a level that is also observed in cultures grown on galactose as the sole carbon source (Figure 4 E). Notably, during the transition towards galactose utilization, transient sub-populations of cells growing on either glucose or galactose were observed (Figure 4 F), eventually leading to full activation of the P_{BL1694} promoter in all cells of the culture (Figure 4 E). These results demonstrate a community-wide adaptation to the utilization of galactose.

Taken together, these results strongly suggest that in *B. longum* glucose mediated catabolite repression is exerted solely on specific sugars such as galactose while not affecting other sugars such as arabinose. The physiological consequences of this differential glucose regulation led to sequential consumption of glucose and galactose but to co-consumption of glucose and arabinose.

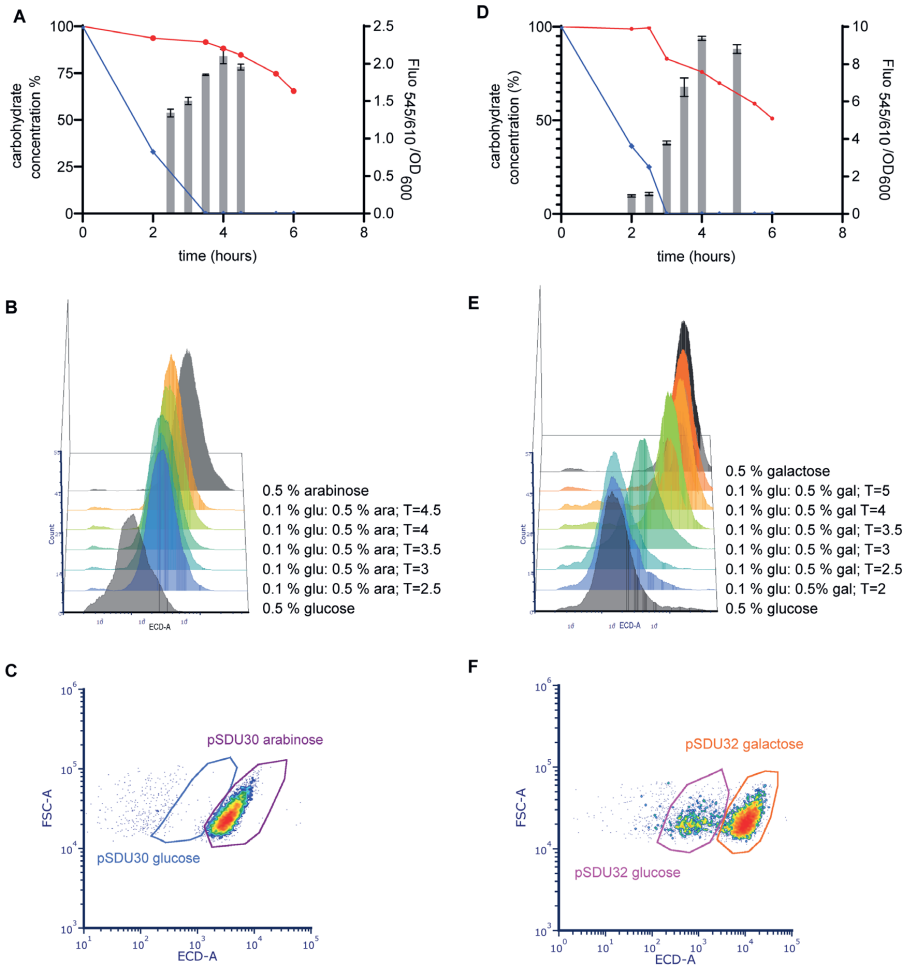


Figure 4: Behavior of *B. longum* NCC 2705 harboring the pSDU30 (P_{BL1359}-mCherry) plasmid cultured on 0.1 % glucose and 0.5 % arabinose (panels A, B & C) or harboring the pSDU32 (P_{BL1694}-mCherry) plasmid cultured on 0.1 % glucose and 0.5 % galactose (panels D, E & F). Above panels show the evolution of the fluorescence signal emitted around the glucose depletion timepoints (grey bars) and the carbohydrate consumption patterns (panel A: glucose (blue) and arabinose (red); panel D glucose (blue) and galactose (red)). Middle panels present the evolution of the populations measured by flow-cytometry around the glucose depletion timepoint (colored), comparing them to glucose and arabinose (grey) (panel B) or glucose and galactose grown cultures (panel E). Below panels represent the population distribution 1h after glucose depletion of both pSDU30 harboring cells (T=4.5, panel C) and pSDU32 harboring cells (T=4, panel F).

Conclusions & discussion

Bifidobacteria are saccharolytic bacteria, and their capacity to metabolize a large range of carbohydrate renders them particularly adapted to the complex gut ecosystem. Detailed study of the fine regulation mechanisms controlling their carbohydrate metabolism requires single cell level molecular tools. Fluorescent protein-based reporters represent therefore an interesting alternative to the commonly used glycan-hydrolase based reporter systems. Fluorescent proteins have been previously expressed in bifidobacteria to enable their real-time detection in *in vitro* or *in vivo* systems [211,212], but to the best of our knowledge, they have not yet been applied in this genus to study the detailed carbohydrate regulation of selected promoters.

In the present study, we evaluated whether fluorescent reporter-proteins that do or do not require oxygen to mature (mCherry and Evoglow-Pp1, respectively) could be used to measure the activity of specific promoters with better precision than the commonly used GusA reporter. Fluorescent proteins that do not require aerobic conditions for their maturation would be most compatible with the anaerobic lifestyle of bifidobacteria. Out of the 3 tested anaerobic fluorescent proteins, Pp1 was identified as the most promising candidate. However, a major drawback of the Pp1 reporter system is that wild-type bifidobacterial strains, including *B. longum* NCC 2705, exhibit a relatively high level of autofluorescence at the excitation and emission wavelengths of this reporter, which decreases the sensitivity and dynamic range of the Pp1-based promoter-reporter system in bifidobacterial strains. The mCherry reporter displayed the most attractive characteristics as it led to a high intensity signal combined with very low or even barely detectable autofluorescence in *B. longum* NCC 2705 as well as other *Bifidobacterium* spp. The requirement for oxygen exposure for mCherry maturation restricts the use of this reporter protein to post-cultivation analysis, which in general would be considered a disadvantage. Nevertheless, by including an appropriate incubation period in the presence of oxygen, we successfully used this protein as reporter to accurately measure promoter activities in anaerobically grown *B. longum* NCC 2705. The design of an anaerobic fluorescent marker with excitation and emission wavelengths similar to those of mCherry would facilitate *in situ* promoter activity evaluation in growing *B. longum* NCC 2705. Interestingly, the biarsenical-tetracysteine tag system was successfully employed for imaging and detection of the anaerobic gut bacterium *Bacteroides thetaiotaomicron* [213]. This system can covalently

react with the green fluorescein (FlAsH) or the red resofurin (ReAsH)[214] and may thereby represent an interesting alternative to mCherry, although the previously reported non-specific ligand binding could present challenges for the use of this system as a promoter-reporter [215].

Nevertheless, we took advantage of the single cell resolution of the mCherry based reporter system to decipher the carbohydrate regulation patterns of the P_{BL1359} and P_{BL1694} promoters predicted to be respectively regulated by AraQ and AraU, two regulatory proteins implicated in arabinose utilization [181,182,198]. The Lac-I type regulator AraQ was additionally proposed to act as global transcriptional regulator in bifidobacteria [181,198], thus potentially representing a functional homologue of the typical Gram-positive CcpA which appears to be absent in *Bifidobacterium* strains. The mechanism by which AraQ may exert its global regulatory role is however likely different to that of the negative regulator CcpA, as AraQ acts both as an activator and a repressor [181,198]. We demonstrated that P_{BL1359}, located upstream of the aldose-1-epimerase encoding gene (BL1359; *galM*) was mildly activated when *B. longum* NCC 2705 was grown on various sugars, including fructose, galactose and ribose, relative to the activity level observed when cells were grown on glucose. These results are in line with the proposed glucose mediated catabolite repression that was recently reported for the regulation of serpin expression in the same strain [140]. Furthermore, the enolase encoding gene *BL1022* predicted to be regulated by AraQ was shown to be differentially expressed in glucose, xylose and galactose, further supporting the role of AraQ in catabolite repression in *B. longum* NCC 2705 [44]. Notably, the P_{BL1359} activity was most prominently induced in cells growing on arabinose as a carbon source, which is in agreement with the prediction that this promoter is regulated by AraQ [180]. The mCherry based reporter allowed to demonstrate that the arabinose-mediated induction of the P_{BL1359} promoter was to a large extend also occurring in the presence of glucose in the whole cell population, which correlates with the observed co-consumption of these two sugars by *B. longum* NCC 2705. These findings imply that the proposed role of AraQ in glucose-mediated catabolite repression is partially alleviated in presence of arabinose, an intriguing observation that warrants further investigation.

The carbon source regulation characteristics of P_{BL1359} were fundamentally different from those observed for P_{BL1694}. The latter promoter is located upstream of an operon

that is predicted to encode a multi sugar ABC transport system previously shown to be induced when cells are grown on short-chain fructo-oligosaccharides [19]. This operon was recently annotated as *araFDH* and proposed to play a role in arabinose transport. The ROK family repressor protein AraU encoded by the gene *BL1693* was hypothesized to control its own expression, but also that of the neighboring *araFDH* operon through regulation of the P_{BL1694} promoter [180,198]. Our data clearly establish that P_{BL1694} was activated during growth on several hexoses like fructose and galactose in comparison to growth on glucose. It was however barely activated by growth on xylose and arabinose, which does not support the previously suggested role of the *araFDH* encoded ABC transport system in arabinose transport [198]. In addition, using the mCherry reporter system, we could demonstrate that the galactose mediated activation of P_{BL1694} is inhibited in the entire cell population in presence of glucose, which is in agreement with the observation that galactose is not consumed when glucose is also present. Based on these observations, we propose that the ROK family regulator AraU and its downstream ABC transporter (*araFDH*) play a role in galactose (and potentially fructose) transport in *B. longum* NCC 2705. The activation pattern of P_{BL1694} shares several similarities to the arabinose inducible P_{BAD} promoter that has been used to construct widely used sets of inducible expression vectors in *E. coli* [216,217]. Therefore, the P_{BL1694} promoter could represent an interesting option for the construction of a galactose-inducible gene expression system for *B. longum* NCC 2705 and possibly other bifidobacteria in which the regulation of the AraFDH transporter is conserved.

Taken together, we demonstrated that the mCherry promoter-reporter system could recapitulate the relative activity levels of promoters previously characterized using the more traditional β-glucoronidase reporter [203]. In addition, the mCherry reporter allowed to study promoter activation at single cell level, which allowed us to unravel the population wide carbohydrate-activation profiles of two promoters that were previously suggested to play an important role in arabinose consumption (as well as other sugars) by *B. longum* NCC 2705 [198]. We established that *B. longum* NCC 2705 behaves differently depending on the carbohydrates it is exposed to, displaying distinct transition phenotypes when switching from glucose to arabinose or to galactose, where glucose and arabinose appeared to be co-metabolized, glucose and galactose were used sequentially. Interestingly, the time required to reach glucose depletion was

extended when this sugar was co-metabolized with arabinose, suggesting that co-metabolization reduces the utilization rate for both sugars. Finally, during the transition from glucose to galactose utilization we observed the transient coexistence of two sub-populations that probably grow on either glucose or galactose. Nevertheless, in *B. longum* NCC 2705 the whole population shifted finally to galactose utilization. This population-wide adaptation in NCC 2705 contrasts with previous *L. lact*is observations where only a specialized sub-population is at the origin of the glucose to galactose metabolic shift [218].

Our results illustrate how fluorescent protein promoter-reporters can be used to shed light on the complex regulation of carbohydrate metabolism in *B. longum* NCC 2705. Especially when combined with flow cytometry, these reporters allow the study of (sub-)population adaptation behavior in *B. longum* NCC 2705 residing and growing in environments containing more than a single carbohydrate substrate. In the dynamic environment of the intestinal tract it is highly plausible that bifidobacteria are frequently exposed to different carbohydrate substrates simultaneously, which underpins the relevance of approaches like the one we followed here to unravel how these bacteria coordinate their metabolic versatility and adaptation under such conditions.

Materials and Methods

Plasmid constructs. Intensity of different aerobic and anaerobic fluorescent proteins was tested in *B. longum* NCC 2705. Plasmid pMDY23 [203] and its derivatives encoding GFP (pVG-GFP), YFP (pVG-YFP), CFP (pVG-CFP) and mCherry (pVG-mCherry) under the control of the strong *B. bifidum* P_{gap} promoter were generously provided by Dr. Christian Riedel (Ulm University, Germany) (Table S 1) [212]. The sequences coding for three different cyan-green anaerobic fluorescent proteins (Bs1, Bs2 and Pp1) were obtained from Evocatal GmbH (Düsseldorf, Germany). Codon optimized (for expression in bifidobacteria) variants of these genes were synthesized (Thermo Fisher Scientific/Geneart GmbH, Regensburg, Germany) and cloned under the control of the P_{gap} promoter by replacing the mCherry encoding region (encompassed between XhoI and SacII restriction sites) of pVG-mCherry, yielding pSDU01 (P_{gap}-Bs1), pSDU02 (P_{gap}-Bs2) and pSDU03 (P_{gap}-Pp1) (Table S 1).

The Pp1 and mCherry fluorescent proteins were subcloned under control of the previously described raffinose-inducible promoter preceding *BL1518* in *B. longum* NCC 2705 (P_{BL1518}; genome coordinates 1692388 to 1692679) [203]. The P_{BL1518} fragments harboring different restriction sites at each end were synthesized and cloned in front of mCherry and Pp1 encoding genes (see pVG-mCherry plasmid map, Figure S 1) (Thermo Fisher Scientific/Geneart GmbH). For mCherry this was achieved by replacing the BglII-MscI P_{gap} promoter fragment in pVG-mCherry by the similarly digested P_{BL1518} fragment, yielding pSDU12 (P_{BL1518}-mCherry). Similarly, P_{BL1518} was used to replace the P_{gap} region in pSDU03 using the BglII-BclI cloning sites, generating pSDU15 (P_{BL1518}-Pp1) (Thermo Fisher Scientific/Geneart GmbH) (Table S 1). Two plasmids encoding β-glucoronidase under the control of the P_{gap} (P_{gap}-*gusA*; pMDY23 [212]) and the raffinose inducible P_{BL1518} (P_{BL1518}-*gusA*; pGUSC [203]) were used as controls (Table S 1).

The promoter preceding the aldose-1-epimerase encoding gene (P_{BL1359} ; genome coordinates 262095 to 261970; reversed sequence) and the promoter upfront of the *BL1694* gene encoding the proposed pentose binding protein (P_{BL1694} ; genome coordinates 2114397 to 2114649) (Figure S 5) were synthesized and subcloned in the pVG-mCherry plasmid replacing the BglII-MscI P_{gap} promoter fragment in pVG-mCherry, resulting in the pSDU30 (P_{BL1359}-mCherry) and pSDU32 (P_{BL1694}-mCherry)

plasmids, respectively (Thermo Fisher Scientific/Geneart GmbH). The orientation of both sequences used for the cloning was carefully selected to decipher the P_{BL1359} and P_{BL1694} promoter activities. All plasmids carried a spectinomycin resistance gene as selective marker.

Strain cultivation and transformation. All strains were retrieved from the Nestle Culture Collection (NCC, Nestlé Research, Lausanne, Switzerland) and were routinely cultured in MRS medium supplemented with 0.05% cysteine-HCl (MRSc), under anaerobic conditions at 37°C, in a closed jar containing an AnaeroGen sachet (ThermoFischer Scientific GmbH). *B. longum* NCC 2705 and *B. breve* ATCC 15700 (T) electrocompetent cells were prepared using the method previously described by Serafini et al. [199]. Strains were grown to stationary phase (overnight incubation) in fresh MRSc supplemented with 16% (w v^{-1}) of short chain fructo-oligosaccharides (scFOS) (Actilight®; Beneo-Orafti, Oreye, Belgium) at 37°C in anaerobiosis. This preculture was diluted 1:10 in 200 ml of fresh MRSc broth and 16% scFOS, and cultivated at 37 °C until an OD_{600} nm of 0.6–0.7 was reached. Cultures were cooled on ice and cells were harvested by centrifugation (2500 g for 10 min), then washed twice with 20 ml of ice-cold washing buffer (16 % FOS in a 1 mM citrate buffer, pH 6.0). Finally, cells were resuspended in 800 µl (1/250 of the original culture volume) of the same ice-cold buffer. Aliquots of 80 µl were frozen in liquid nitrogen and stored at -80 °C until use. Cells were thawed on ice and plasmid DNA (200 ng) was transformed by electroporation in a Nucleofector device and using the AA-035 protocol settled in the apparatus (RUWAG Handels AG, Bettlach, Switzerland). Electroporated cells were recovered in 1 ml of fresh MRSc and incubated for 2-3h in anaerobiosis at 37°C before plating. Transformant colonies were obtained after 48h of growth at 37°C in anaerobiosis on MRSc agar containing 200 µg ml^{-1} of spectinomycin (MRScs). Transformants were grown anaerobically overnight at 37° in MRScs and stored at -80°C after addition of 20% glycerol (v v^{-1}).

β-glucoronidase activity measurement. Overnight anaerobically grown cultures of *B. longum* NCC 2705 containing the plasmids carrying the β-glucoronidase encoding gene (*gusA*) under the control of the different promoters were subcultured in 30 ml of fresh MRSc based medium lacking a carbon source (MRSc-C) (10 g l^{-1} of bacto proteose peptone n°3, 5 g l^{-1} bacto yeast extract, 1 g l^{-1} Tween 80 [all from Chemie Brunschwig, Buchs, Switzerland], 2 g l^{-1} di-ammonium hydrogen citrate, 5 g l^{-1} sodium

acetate, 0.1 g l^{-1} magnesium sulphate, 0.05 g l^{-1} manganese sulfate, 2 g l^{-1} di-sodium phosphate, 0.5 g l^{-1} cysteine [all from Sigma-Aldrich Chemie GmbH]) to which 2% of glucose or raffinose was added. Growth experiments were performed in biological triplicates, at 37°C under anaerobiosis and until mid-logarithmic phase (OD_{600} ~0.6). All media contained 200 µg ml^{-1} of spectinomycin.

Cells were harvested from mid-exponential cultures by centrifugation (2500 g for 10 min at 4°C) and concentrated to reach 12 OD_{600} units per ml in the previously described GusA-assay buffer [212] consisting of 50 mM Na_2HPO_4 (pH 7), 1 mM EDTA, 0.1% Triton X-100 and 5 mM dithiothreitol (DTT). Cells were mechanically disrupted using glass beads, for 3 times 1 minute at 4 m s^{-1} using a FastPrep-24 (MP Biomedicals LLC, Irvine, USA). Cell debris were removed by centrifugation (3300 g, 5 min, 4°C) and 100 µl of the cell lysate was transferred to 96 well plates before being mixed with 100µl of 15 mM 4-Nitrophenyl β-D-glucuronide (PNP-GLUC; Sigma-Aldrich Chemie GmbH) dissolved in assay buffer. Plates were incubated at 37°C and absorbance was measured at 405 nm every 5 minutes for 30 minutes in a Varioskan spectrophotometer (Thermo Fisher Scientific). Enzyme activities were expressed as nitrophenol nanomoles released per minute.

Promoter activity measured through fluorescence level determination. Overnight cultures of *B. longum* NCC 2705 harboring the plasmids encoding fluorescent proteins under control of the different promoters were cultured as previously described in MRSc-C to which 2% of the respective sugars (glucose, raffinose, fructose, galactose, ribose, arabinose or xylose) was added. Precultures were performed in the same medium composition to preadapt the bacterial cells to the respective substrates.

To determine cyan-blue Evoglow-Bs1, Bs2 and Pp1 fluorescence levels, cells were washed twice and resuspended in 250mM phosphate buffer at pH 7.0. The OD_{600} of the suspensions as well as fluorescence were measured in a Varioskan spectrophotometer (Thermo Fisher Scientific) without any further incubation. Excitation and emission wavelengths of 450 and 495 nm, respectively, were used for all variants.

To determine fluorescence levels of proteins requiring oxygen to mature, cells were centrifuged (3500 g for 2 min), washed twice and resuspended in 250mM phosphate buffer at pH 7.0 containing the bacteriostatic antibiotic Chlortetracycline (Sigma-

Aaldrich chemie GmbH) at 30 µg ml^{-1} to stop protein synthesis. Cells were then incubated at room temperature (in presence of oxygen) for 90 min, and OD_{600} and fluorescence levels were measured in a Varioskan spectrophotometer (Thermo Fisher Scientific). Excitation and emission wavelengths were the following: 470 and 625 nm for GFP, 436 and 480 nm for CFP, 500 and 535 nm for YFP, and 545 and 610 nm for mCherry, respectively.

Statistical analysis. Statistical analysis and graphical representations of the different promoter activation datasets was performed using GraphPad prism v8.1.1 (GraphPad Software, San Diego, USA). Average and standard deviations are represented, while significance levels were determined using a one-way ANOVA, followed by Sidak's multiple comparison tests (α=0.05).

Measurement of population evolution by flow-cytometry. Cells harboring the pSDU30 (P_{BL1359}-mCherry) and pSDU32 (P_{BL1694}-mCherry) plasmids were grown in fresh MRSc-C supplemented with 200 µg/ml spectinomycin and different carbohydrate mixtures (0.1 % glucose:0.5 % arabinose and 0.1% glucose:0.5% galactose, respectively). Ten (10) ml of early exponential phase cultures were aliquoted in smaller volumes (1ml) and incubated at 37°C in anaerobiosis for another 12 hours and collected at regular intervals throughout the exponential phase. Cells were harvested by centrifugation (room temperature, 3000 g, 2 minutes) and resuspended in 250mM phosphate buffer at pH 7.0 containing 30 µg ml^{-1} Chlortetracycline (Sigma-Aldrich Chemie GmbH, Buchs, Switzerland). Population distribution was evaluated by flow cytometry after 90 minutes of aerobic incubation at 37°C, which corresponds to the optimal maturation time of the mCherry protein. Residual glucose levels were determined in these cultures using the MQuant kit (Sigma-Aldrich Chemie GmbH) according to the manufacturer's protocol and carbohydrate consumption was determined by HPLC (see below).

Flow cytometry analyses were performed with a Beckman Coulter Cytoflex S (Beckman Coulter, Brea, US) equipped with four lasers. The mCherry fluorescence was measured in the ECD detector (excitation at 561nm, emission at 610/20 nm). Results were analysed with the Beckman-Coulter CytExpert Acquisition software version 2.3. All data were exported in FCS 3.0 format and analysed offline with the De Novo FCSExpress software version 7.01.

Carbohydrate consumption measured by HPLC. Glucose, galactose and arabinose levels were quantified in filtered sterile culture supernatants by High Performance Liquid Chromatography (HPLC, Agilent Technologies AG, Basel, Switzerland), using an Aminex HPX-87H columm (Bio-Rad Laboratories AG, Cressier, Switzerland). 5 µl of samples were injected, using a 5 mM H_2SO_4 solution as eluent, with a flow rate of 0.6 ml/min (run time 25 min). Chromatogram was obtained using a refractive index detector, from which peak areas and corresponding concentrations were calculated.

Supplementary data

Table S 1: List of plasmids used in this work

Material	Description	Origin
pVG-YFP	YFP fluorescent protein under the control of *B. bifidum* S17 P_{gap}	Derived from pMDY23, Grimm et al.
pVG-CFP	CFP fluorescent protein under the control of *B. bifidum* S17 P_{gap}	Derived from pMDY23, Grimm et al.
pVG-GFP	GFP fluorescent protein under the control of *B. bifidum* S17 P_{gap}	Derived from pMDY23, Grimm et al.
pVG-mCherry	mCherry fluorescent protein under the control of *B. bifidum* S17 P_{gap}	Derived from pMDY23, Grimm et al.
pSDU01	Bs1 anaerobic fluorescent protein under the control of *B. bifidum* S17 P_{gap}	Derived from pVG-mCherry, this work.
pSDU02	Bs2 anaerobic fluorescent protein under the control of *B. bifidum* S17 P_{gap}	Derived from pVG-mCherry, this work.
pSDU03	Pp1 anaerobic fluorescent protein under the control of *B. bifidum* S17 P_{gap}	Derived from pVG-mCherry, this work.
pMDY23-pGap	β-glucoronidase reporter enzyme under the control of *B. bifidum* S17 P_{gap}	Derived from pMDY23, Grimm et al.
pGusC	β-glucoronidase reporter enzyme under the control of *B. longum* NCC 2705 P_{BL1518}	Derived from pMDY23, Klijn et al.
pSDU12	mCherry fluorescent protein under the control of *B. longum* NCC 2705 P_{BL1518}	Derived from pVG-mCherry, this work.
pSDU15	Pp1 anaerobic fluorescent protein under the control of *B. longum* NCC 2705 P_{BL1518}	Derived from pSDU03, this work.
pSDU30	mCherry fluorescent protein under the control of *B. longum* NCC 2705 P_{BL1359}	Derived from pVG-mCherry, this work
pSDU32	mCherry fluorescent protein under the control of *B. longum* NCC 2705 P_{BL1694}	Derived from pVG-mCherry, this work

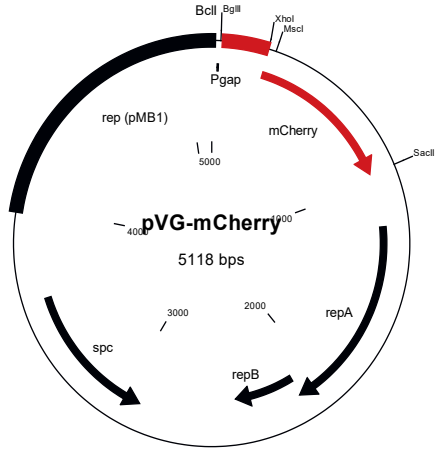


Figure S 1: Map of the pVG-mCherry plasmid, highlighting the restriction sites used for cloning in this work and containing the following elements: pMB1 replicon for replication in *E. coli* (black region); *repA* and *repB* for replication in bifidobacteria and the spectinomycin resistance gene (black arrows); P_{gap} promoter (red region); genes encoding the fluorescent proteins mCherry (red arrow).

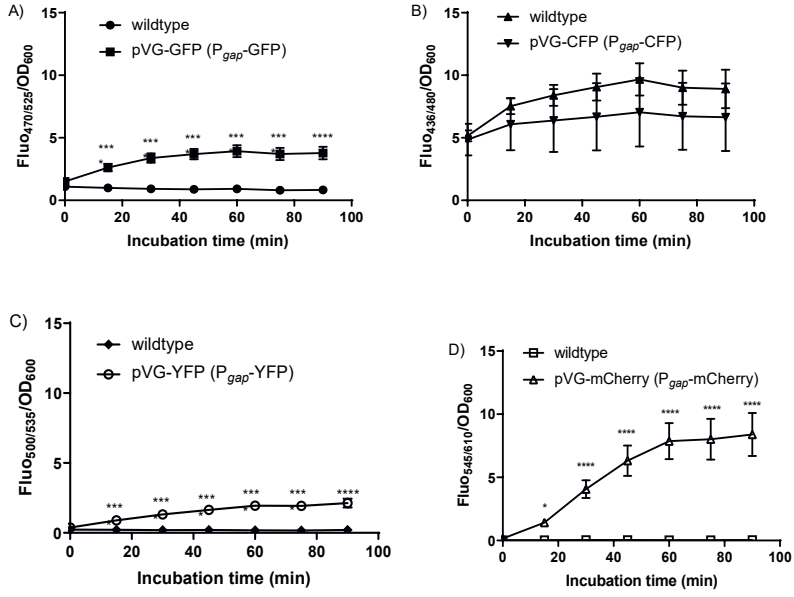


Figure S 2: Fluorescence signals of MRSc stationary grown wildtype *B. longum* NCC 2705 and its recombinant derivatives harboring pVG-GFP (A), pVG-CFP (B), pVG-YFP (C) and pVG-mCherry (D) measured at the respective excitation and emission wavelengths of each individual reporter proteins. Data represent averages and standard deviations of biological triplicates and were determined using technical duplicates. P-values were calculated using one-way ANOVA, followed by a Sidak's multiple comparison test. **** p-value < 0.0001 as compared to control (wildtype).

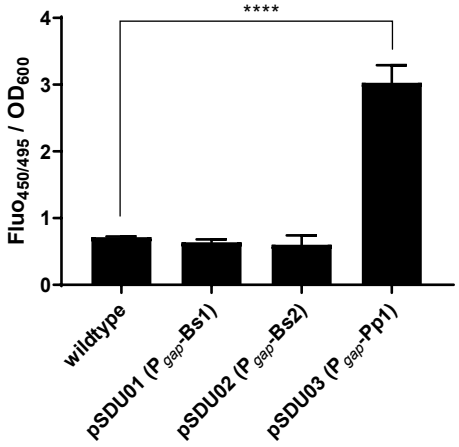


Figure S 3: Fluorescence signal of stationary grown wildtype *B. longum* NCC 2705 and its pSDU01, pSDU02 or pSDU03 harboring transformants measured at excitation and emission wavelengths of 450 and 495 nm, respectively. Values represent average of biological triplicates. Significant differences between the background fluorescence of the wild-type and recombinant strains were calculated using one-way ANOVA, followed by a Sidak's multiple comparison test (**** p-value < 0.0001).

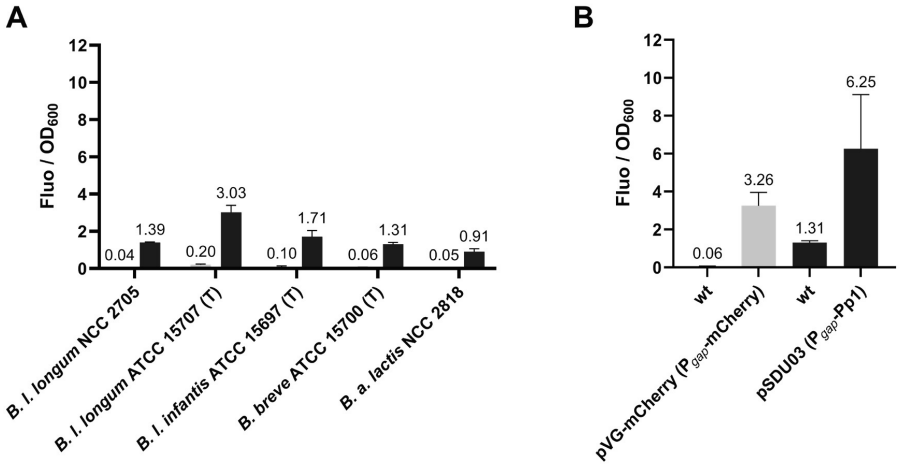


Figure S 4: Comparison of fluorescence emitted by mCherry and Pp1 reporter proteins in different bifidobacterial hosts. Fluorescence signal of MRSc stationary grown wildtype *B. longum* NCC 2705, *B. longum* ATCC 15707 (T), *B. longum* subsp. *infantis* ATCC 15697 (T), *B. breve* ATCC 15700 (T) and *B. animalis* subsp. *lactis* NCC 2818 measured at excitation and emission wavelengths of 545 and 610 nm (light grey bars), as well as 450 and 495 nm (dark grey bars), respectively (panel A). Fluorescence signal emitted at the same wavelengths by wildtype and P_{gap} recombinant derivatives of *B .breve* ATCC 15700 (T) (panel B). Values represent average of biological duplicates.

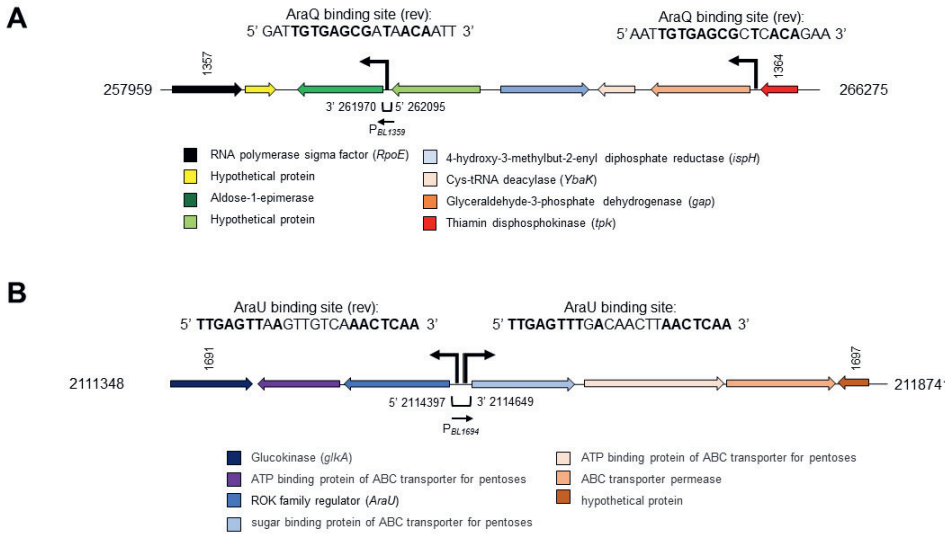


Figure S 5: Operon organization of the two regions encompassing AraQ (Panel A) and AraU (Panel B) targeted cloned promoters (P_{BL1359} and P_{BL1694}). Genes and orientations are indicated by colored arrows. The promoter regions and their directions are indicated in respective intergenic regions together with the predicted RegPrecise binding sites. Conserved binding motifs are highlighted in bold.

CHAPTER 6

6

The pleiotropic effects of carbohydrate-mediated growth rate modifications in *Bifidobacterium longum* NCC 2705

S. Duboux* [1,2] | S. Pruvost [1] | C. Joyce [1] | B. Bogicevic[1] | J.A. Muller [1] | A. Mercenier [2] | M. Kleerebezem [2]

[1] Nestlé Research, Route du Jorat 57, CH 1000 Lausanne 26, Switzerland
[2] Host-Microbe Interactomics Group, Wageningen University & Research, De Elst 1, 6708WD Wageningen, The Netherlands

Abstract

Bifidobacteria are saccharolytic bacteria that are able to metabolize a relatively large range of carbohydrates through their unique central carbon metabolism known as the “bifid-shunt”. Carbohydrates have been shown to modulate the growth rate of bifidobacteria, but unlike for other genera (e.g., *E. coli* or *L. lactis*), the impact it may have on the overall physiology of the bacteria has not been studied in detail to date. Using glucose and galactose as model substrates in *Bifidobacterium longum* NCC 2705, we confirmed that the strain displayed fast and slow growth rates when grown on those carbohydrates, respectively. We demonstrate that these differential growth conditions are accompanied by global transcriptional changes and adjustments of central carbon fluxes. In addition, when grown on galactose, NCC 2705 cells were significantly smaller, exhibited an expanded capacity to import and metabolize different sugars and displayed an increased acid-stress-resistance, a phenotypic signature associated with generalized fitness. We predict that part of the observed adaptation is regulated by the previously described bifidobacterial global transcriptional regulator AraQ, which we propose to reflect a catabolite-repression like response in *B. longum*. With this manuscript, we demonstrate that not only growth-rate but also various physiological characteristics of *B. longum* NCC 2705 are responsive to the carbon source used for growth, which is relevant in the context of its lifestyle in the human infant gut where galactose-containing oligosaccharides are prominent.

Introduction

Bifidobacteria are important members of the human gut microbiota that particularly strive in early infancy [7]. Their presence in that stage of life has been proposed to play a critical role in the maturation of the newborn's immune system as several studies have linked low abundance of members of this genus in the infant gut to later in life immune disorders such as atopy [38] or asthma [39]. Early-life *Bifidobacterium* dominance has also been proposed to be linked with reduced risk of acquisition of antimicrobial resistances [41] or obesity [40] later in life.

Reflecting their adaptation to the human gut environment, bifidobacteria harbor an extensive carbohydrate-active enzyme repertoire that supports their capacity to effectively breakdown and metabolize a large range of carbohydrate substrates [16,219]. In particular, the predominance of *Bifidobacterium* in early infancy has been attributed to their capacity to grow on Human Milk Oligosaccharides (HMO) that besides lactose dominate the available sugars found in human milk [169]. Strains belonging to *B. longum* subsp. *infantis* have been shown to be particularly adapted to import and hydrolyze a range of fucosylated or sialylated HMOs [8,10,220], while the capacity to metabolize the neutral Lacto-N-Tetraose (LNT) is considered to be broadly conserved in all subspecies of *B. longum* [219,221]. In members of the *Bifidobacterium* genus, all carbohydrates are proposed to be processed through a relatively unique central carbon metabolic pathway, the so-called "bifid-shunt", which involves the bifunctional xylulose 5-phosphate/fructose-6-phosphate phosphoketolase enzyme (F6PPK) [21]. As an example, glucose, fructose, mannose, xylose, ribose and galactose were all shown to be catabolized via the "bifid-shunt" in *B. longum* NCC 2705 [44].

Carbohydrate metabolism is the predominant source of ATP for bifidobacteria [51], and carbon import and metabolic fluxes should be tightly controlled to allow nutrient adaptation and maintenance of fitness under fluctuating environmental conditions. In Gram-positive bacteria, catabolite repression (often mediated by glucose or lactose) has been shown to control a wide array of processes, including the repression of expression of import functions for alternative sugar sources, and regulation of central carbon metabolism flux [176]. The same regulation process has not been extensively investigated in bifidobacterial strains, although several manifestations of glucose-mediated catabolite repression have been observed, suggesting its existence in this

genus as well. For example, glucose was demonstrated in *B. longum* to repress the production of *Frk*, a fructokinase required for the catabolism of fructose [177]. In *B. lactis* glucose was shown to repress the expression of sucrose utilization genes [178]. We recently demonstrated that in *B. longum* NCC 2705, growth on glucose inhibited production of the serine protease inhibitor (serpin) which was relieved when the strain was grown on different sugars [140]. In most gram-positive bacteria, catabolite repression is controlled by a protein complex of histidine phosphocarrier protein (HPr) and catabolite control protein A (CcpA), a LacI-type regulator that can bind to specific DNA motifs (so-called *cre*-elements) and inhibit or activate transcription of downstream genes [179,222]. Of note, recognizable orthologues of the CcpA protein are absent in bifidobacteria and typical genome-wide catabolite repression mechanisms have not been described in detail for the members of this genus. Nevertheless, two LacI-type transcriptional regulators, AraQ and MalR1, were demonstrated to control a large set of genes spread over the genome of *B. breve* UCC 2003. Both proteins were predicted to regulate a range of transcriptional factors in this strain, to repress transcription of a number of genes related to sugar import, and to activate transcription of genes related to central carbon metabolism including genes involved in the ”bifid-shunt” [181,182]. Therefore, AraQ and MalR1 could represent the central regulators of the bifidobacterial variant of the canonical CcpA-mediated catabolite regulation mechanism.

In most prokaryotes, consumption of glucose is coupled to fast growth, which at least in part results from the tight catabolite repression control [176]. Relief of catabolite repression enables bacteria to consume a larger variety of carbon sources, which often results in a reduction of growth rate [223]. These slow-growing bacterial cells are not only better prepared for the utilization of alternative carbon substrates but are also displaying increased resistance to stress [224-226]. These effects were predominantly demonstrated using cells grown at different rates in carbon-limited chemostats, but recent evidences suggest that the same outcome can be obtained by varying the carbon source used for growth [227]. It has been shown previously that bifidobacteria display different growth rates depending on the available carbohydrate substrate. For example, most tested bifidobacterial strains displayed a reduced growth rate on galacto-oligosaccharides (GOS) relative to growth on glucose [227], while the monosaccharide galactose induced a substantial growth rate reduction in *B. longum*

and *B. adolescentis* strains [228]. However, the overall physiological impact of sugar-dependent growth rate-modifications and the possible involvement of catabolite-repression like control mechanisms remain unresolved in bifidobacteria.

The present study aimed to investigate whether carbohydrate induced growth rate variation may influence the physiology of *B. longum* NCC 2705. Glucose and galactose were selected not only for their capacity to modulate the growth rate of the strain, but also because they are the common carbon sources in bifidobacterial growth media [229] as well as important substrates for bifidobacteria in the infant colon as they are core-constituents of HMOs [169]. We show that growth on glucose and galactose led to distinct metabolic and physiological changes and that this was associated with substantial genome-wide transcriptional modifications. Overall, this study provides a further demonstration of the pleiotropic impact of specific carbohydrate substrates on the physiology of this saccharolytic bacterium, which support the existence of a a catabolite-repression like mechanism in *B. longum* NCC 2705.

Results

The distinct growth attributes of *B. longum* NCC 2705 cultured on glucose or galactose as sole carbon source. *B. longum* NCC 2705 was grown in bioreactors, using MRS-based medium supplemented with glucose or galactose as sole carbon source. The maximal growth rate (μ) on glucose supplemented medium was 0.67 h^{-1}, whereas maximal growth on galactose medium was approximately 2-fold slower (μ=0.35 h^{-1}). Estimated doubling time was of 1.05 h when the strain was grown on glucose, while it was twice slower (2.06 h) in the galactose condition. The observed growth rate was reflected by the difference in acidification rates, where glucose and galactose growing cells displayed maximal acidification rates of -0.57 and -0.39 h^{-1}, respectively (Table 1). The maximum number of culturable cells (CFU/ml) in the galactose grown cultures was substantially higher (7.8E8 CFU/ml) than that obtained in the glucose grown culture (2.8E8 CFU/ml). Additionally, a drastic decline in culturability (CFU/ml) was observed between 12 and 20 hours of cultivation on glucose supplemented medium (Figure 1 A), which coincides with the entry into the stationary phase of growth. In contrast, a similar decrease occurred between 30 and 40 hours for the galactose grown culture (Figure 1 B), which may reflect the extended transition period from the exponential to the stationary phase observed for the galactose cultivated cells. Strikingly, and despite the similar final pH reached (4.5 in both conditions), the final optical density obtained in the glucose grown culture (OD_{600} = 5.6) was higher as compared to the galactose grown one (OD_{600} = 3.9). Taken together, these data demonstrated that there was approximatively a 4 times higher number of culturable cells per OD in the galactose grown culture as compared to the glucose grown one, leading to CFU:OD ratios of 2.54E8 and 5.93E7, respectively (Table 1), suggesting a difference in cell size between the two growth conditions. Therefore, we evaluated the respective cell size of glucose and galactose grown *B. longum* NCC 2705 using confocal microscopy and forward scatter values obtained by flow cytometry. These analyses showed that galactose grown cells had a shortened rod-shape (1,57 ± 0,03 μm) as compared to the glucose grown cells (2,47 ± 0,06 μm; Figure 2 A and B), irrespective of the growth phase at which these cells were harvested (Table 1). These cell-length estimations are in agreement with the significantly lower ($p<0.0001$) forward scatter values obtained by flow cytometry analysis of the galactose grown cells relative to their glucose grown counterparts (Figure 2 C), confirming that

galactose grown cells are significantly smaller compared to glucose grown cells. These results highlight pleiotropic modifications of *B. longum* NCC 2705 characteristics induced by growth on glucose or galactose as sole carbon source.

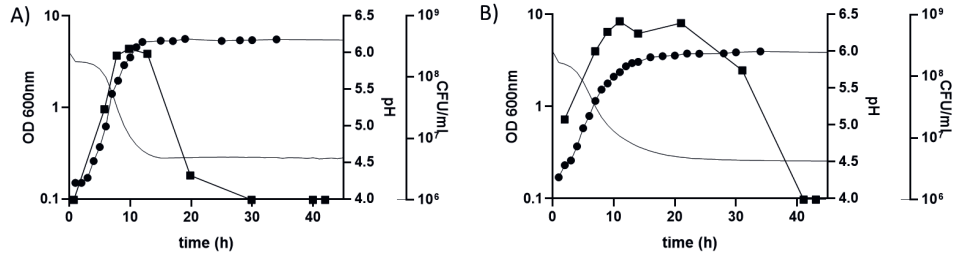


Figure 1: Growth characteristics of *B. longum* NCC 2705 grown in DasGip bioreactors with 1% of glucose (A) or 1% of galactose (B). Optical density measured at 600 nm (OD; black circles), cell concentration measured by Colony Forming Unit (CFUs, black squares) and pH (black line) are depicted for each condition.

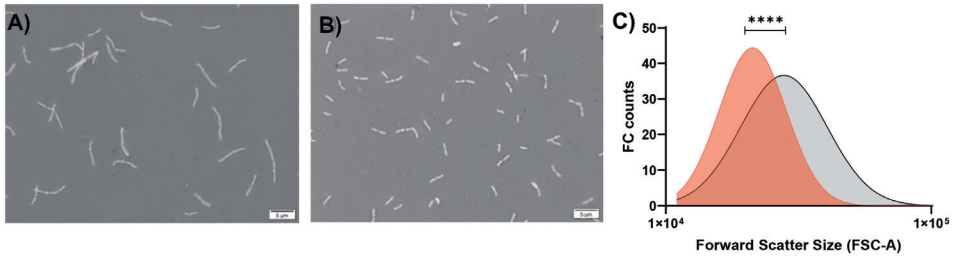


Figure 2: Confocal microscopic pictures of cells of *B. longum* NCC 2705 grown in DasGip bioreactors with 1% of glucose (A) or 1% of galactose (B). Panel C shows the cell size of both cultures (galactose in red, glucose in grey) represented by the forward scatter signal obtained by flow cytometry. Statistical analyses were performed using Welch-t test (**** $p<0.0001$) on modeled population distributions.

Table 1: Growth characteristics of *B. longum* NCC 2705 grown in DasGip bioreactors with 1% of glucose or galactose. Maximal acidification slope was obtained by modeling inversed pH values. Maximal growth rate was retrieved from the modelling of the OD curves. An average CFU:OD ratio was calculated based on all data obtained before the decay in CFU. The estimated doubling time was obtained by modeling the increase of theoretical CFU through time. Average cell length represents the average of all data (early-exponential, mid-exponential and stationary harvested cells) obtained for the different conditions.

	GLUCOSE		GALACTOSE	
	OD-based	**CFU-based**	**OD-based**	**CFU-based**
Max cell density (OD_{600}; CFU/ml)	5,60	2,80E+08	3,90	7,80E+08
Max growth rate μ (1/h)	0.67	-	0.35	-
Estimated doubling time (1/h)	-	1.05	-	2.06
Average CFU:OD	5,93E+07		2,54E+08	
Average cell length ± SEM (μm)	2,47 ± 0,06		1,57 ± 0,03	
Max acidification slope (1/h)	-0.57		-0.39	
Final pH	4,51		4,50	

Glucose or galactose dependent central carbon metabolism fluxes modifications. Transcriptional analysis of the genes related to the carbon import and metabolism (bifid-shunt) revealed that growth on glucose induced the expression of glucose import systems (*Bl1631* [MFS] and *BL1633* [part of PTS]), enzymes involved in Glucose-6-phosphate conversion (*BL1631; 6PGL*), pyruvate and acetate formation (*BL0707* [pgk] and *BL112*4 [aldH]), the *galE1* gene (*BL1644*) encoding the conversion of UDP-glucose to UDP-galactose, as well as the ATP binding function (*BL1692*) of the ABC transport system that was previously annotated to import arabinose [198] (Figure 3). In contrast, growth on galactose induced the expression of a substantially larger number of functions, including the relatively high level induction of carbohydrate transporters (ABC (*BL0189*) and an MFS transport system (*BL0165*), upregulated 3.0 and 3.5 fold, respectively) that we propose to be involved in galactose import. Furthermore, functions related to the catabolism of the imported galactose were upregulated, including the conversion of D-galactose to D-glucose-6-phosphate and D-fructose-6-phosphate (*BL0739* [*ugpA*], *BL1210* [*galK*], *BL1671* [*galE2*], *BL1211* [*galT1*], *BL1643* [*galT2*], *BL1630* [*pgm*] and *BL0279* [*pgi*]). Notably, two genes (*BL1694*, *BL1695*) that encode an ABC system transporter that was previously suggested to be involved in arabinose transport [198] were also induced, similarly to *araA* (*BL0272*) which encodes the conversion of D-arabinose into D-ribulose, suggesting the activation of other sugar catabolism functions unrelated to galactose. Although the gene encoding F6PPK (fructose-6-phosphate-phosphoketolase*, BL0959*) that plays a central role in the bifid-shunt was not differentially expressed, the *tpk* gene (*BL1364*)*,* responsible for the production of thiamin that acts as an essential cofactor for F6PPK was upregulated. The activity of the F6PPK enzyme upon growth on galactose could thus be enhanced, which would be predicted to increase the levels of acetyl-phosphate in the cell. This may be relevant based on the observation that galactose growth leads to high level induction of functions that utilize acetyl-phosphate as a substrate, and convert it to ethanol (*BL1090, adh*; 4.8-fold induced) and formate (*BL0951, pfl*; 9.5-fold induced) (Figure 3).

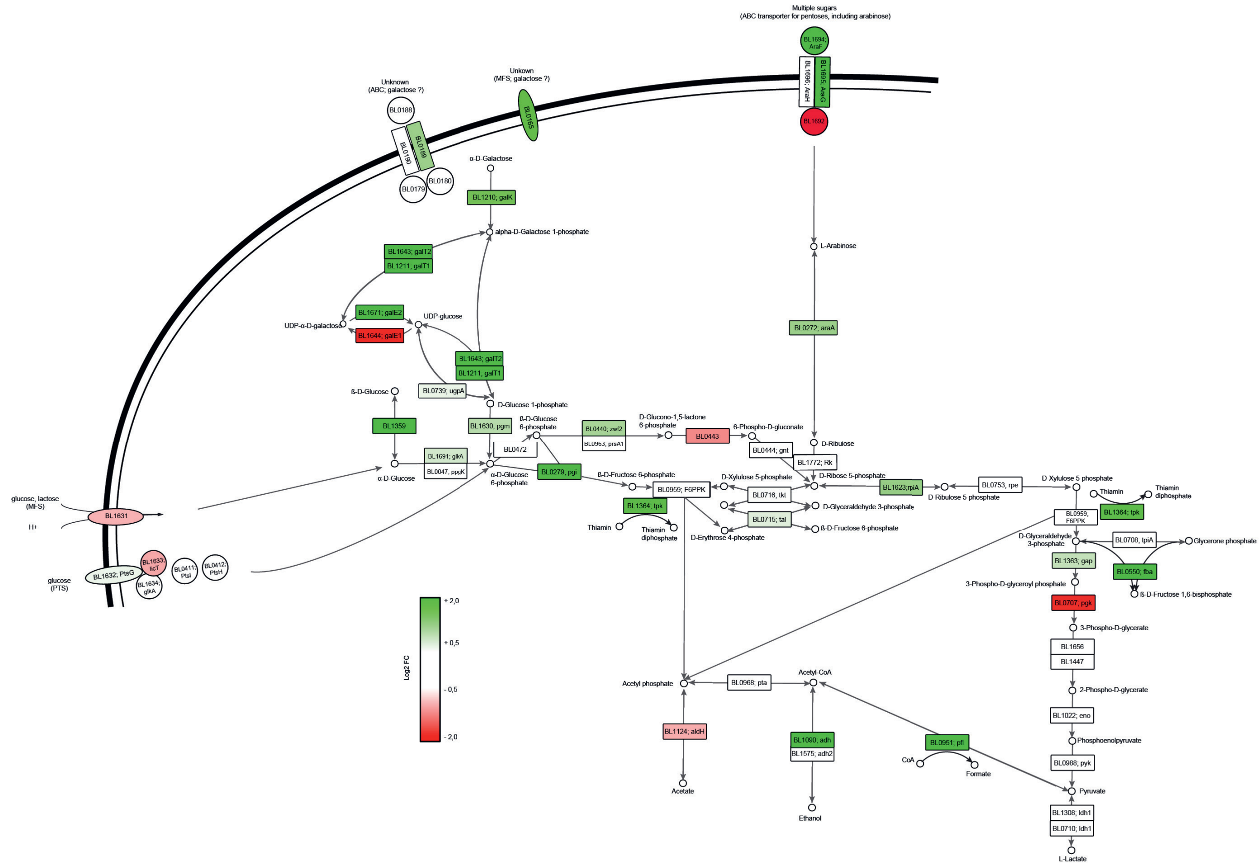


Figure 3: Differential expression of genes related to carbohydrate import and central carbon metabolism of *B. longum* NCC 2705 grown on galactose or glucose. Color scale represent log2 fold change of galactose:glucose (green) or glucose:galactose (red).

6

These transcriptional changes suggest a shift in the central carbon metabolic fluxes dependent on the sugar used for growth. To evaluate this, we quantified fermentation end products of both cultures by Nuclear Magnetic Resonance (NMR) and established that formate and ethanol production increased during growth on galactose (from 0.5 to 6.1 % and 0.9 to 4.8 %, of the overall end-products formed, respectively), which coincided with decreased lactate production (from 45.8 % to 32.6 %), while acetate production appeared to modestly increase (52.8 to 57.3 %) (Figure 4). These findings confirmed that the induction of the pyruvate formate lyase (*pfl*) and alcohol dehydrogenase (*adh*) translates to increased levels of the products that these enzymes produce (formate and ethanol, respectively). The prominent enhancement of the flux towards formate and ethanol (and to a lesser extent acetate) occurs at the expense of the flux towards lactate, involving the *ldh* encoded lactate dehydrogenase that was not differentially expressed. As a consequence, the acetate:lactate ratio shifted from 1.15 in glucose grown cells to 1.76 in galactose grown cells of *B. longum* NCC 2705.

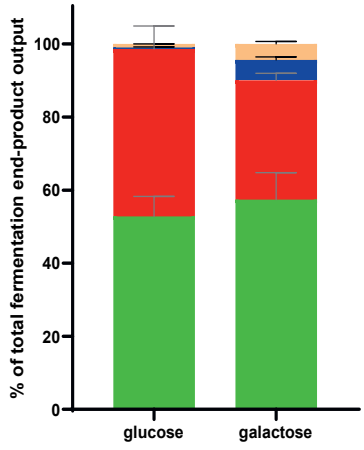


Figure 4: Relative levels of fermentation end-products produced by *B. longum* NCC 2705 grown on glucose or galactose. Acetate (green), lactate (red), ethanol (orange) and formate (blue) levels measured in stationary phase culture supernatants.

Expanded carbohydrate utilization capacity in galactose grown *B. longum* NCC 2705. *B. longum* NCC 2705 grows slower on galactose (Table 1) and upon growth on this carbohydrate, genes implicated in transport and metabolization of another sugar (e.g., arabinose) were induced (Figure 3; see also above). In *L. lactis* and *E. coli*, growth rate reduction was shown to coincide with the production of non-essential proteins implicated in the import and metabolization of alternative sugar sources [230]. Therefore, we evaluated in more detail to what extent genes involved in the transport of carbohydrates (beyond glucose, galactose and arabinose) [19] were differentially expressed in the two conditions studied.

During glucose growth, a gene involved in the transport of mannosylated oligosaccharides (*BL1332*) was slightly but significatively upregulated. In contrast, in galactose growing cells several genes predicted to be involved in the import of various carbohydrates were induced, including the transporters for fructose and FOS (*BL0033*), maltose and arabinose (*BL00141, BL0146*), raffinose (*BL1521, BL1524, BL1526*), xylans and arabinans (*BL0426*), and xylosides (*BL1710*) (Figure 5). Interestingly, differential but opposite regulation was observed for the two substrate-binding proteins associated with an ABC-transport system predicted to import arabinose oligosaccharides (*BL1163* induced 9.4-fold in glucose, *BL1164* induced 2.2-fold in galactose). These data suggest that relative to glucose-grown cells, the slower growing galactose cells are potentially able to utilize a larger set of carbohydrate-substrates, a phenotype adjustment that is referred to as catabolic flexibility [231].

To confirm galactose growth mediated catabolic flexibility, glucose and galactose grown cells were tested in an acidification assay using translationally blocked cells (using 5 µg/ml erythromycin) [232] and providing these with a range of carbohydrates that were previously shown to support NCC 2705 growth [19,140]. As expected, glucose grown cells were best adapted to import and metabolize glucose, while galactose grown cells acidified at the highest rate when provided with galactose. Glucose grown cells could besides glucose also ferment galactose, lactose and GOS, and to a minor extent also LNT. In contrast, galactose grown cells could effectively ferment a substantially larger set of carbohydrates, including not only glucose, galactose, and lactose, but also sucrose, arabinose, scFOS, GOS and LNT, while fermentation activity was also observed when providing fructose, raffinose and maltose albeit at a low rate (Figure 6).

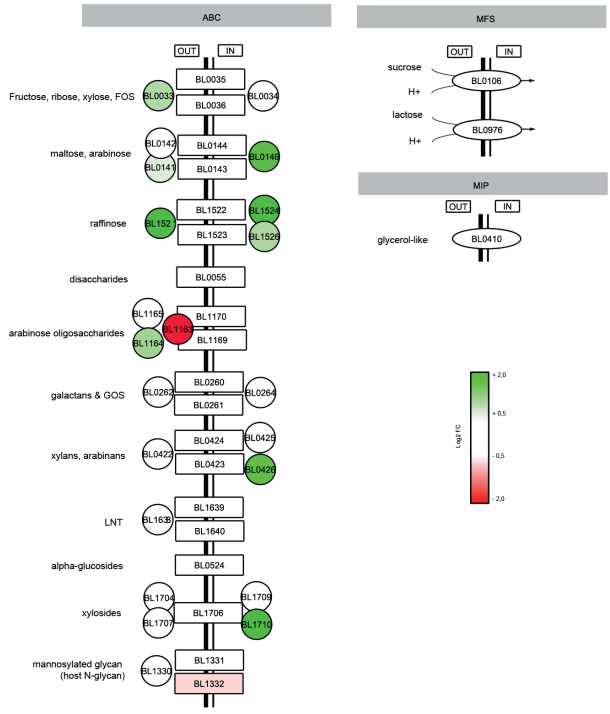


Figure 5: Differential expression of genes related to carbohydrate transport in *B. longum* NCC 2705 grown on galactose or glucose. Fold expression change in galactose growth induced (green) and glucose growth induced genes (red) are displayed in log2 scale.

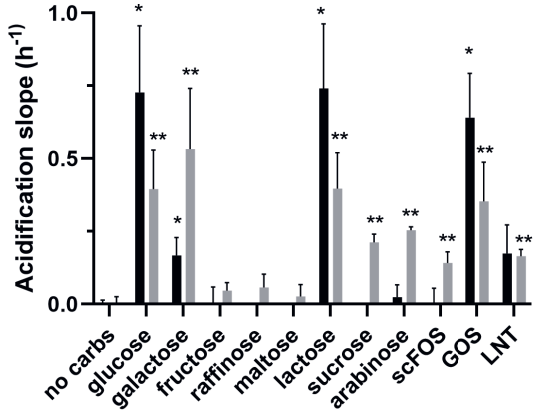


Figure 6: Initial acidification rates of glucose (black bars) or galactose (grey bars) mid-exponentially grown, translationally blocked *B. longum* NCC 2705 cells that were provided different carbohydrates. Values were extracted from the average logarithmic slopes obtained within the first 200 min of incubation after subtracting background acidification (no carbs) levels. Average acidification rates and standard deviation were calculated from biological duplicates measured in duplicate. Statistical analyses were performed using an unpaired multiple t-test comparing all carbohydrate conditions to the "no carbs" control (* q<0.001, ** q<0.0001)

The potential implication or AraQ and MalR1 Lac-I regulators. AraQ and MalR1 have previously been proposed as global bifidobacterial regulators [181]. We therefore determined the potential involvement of these two regulatory proteins in the orchestration of the distinct catabolic flexibility observed when comparing glucose or galactose grown *B. longum* NCC 2705. Out of the 13 genomic loci predicted to be under AraQ transcription control [180], 9 were differentially expressed in our analysis. These loci encode genes that play prominent roles in the observed galactose-mediated adaptations, including the genes involved in the bifid-shunt (*BL0715 [tal], BL0716 [tkt], BL1363 [gap]*) and those involved in the utilization of alternative carbon sources like arabinose (*BL0272 [araA], BL0274 [araB]*), maltose (*BL0141 [malE]),* galactose *(BL1359 [galM*]) and 1,4-alpha-glucans (*BL0999 [glgB]*). Conversely, the gene coding for a biotin-protein ligase (*BL1533*) was downregulated during growth on galactose, similar to the gene encoding AraQ itself (*BL0275*), which could reflect its capacity to both activate and repress transcription and its previously proposed autoregulation (Table 2).

A set of 7 loci are predicted to be regulated by MalR1 in *B. longum* NCC 2705 [180]. Although the MalR1 encoding gene was not differentially expressed, four of its target loci were induced during growth on galactose, i.e., genes involved in maltose transport (*BL0141 [malE*]), channeling of galactose into the bifid-shunt (*BL1630 [pgm]*, *BL0279* [pgi]), and glucose transport and metabolism (*BL1631* [pgm], *BL0279* [pgi]). Notably, two of the MalR1 target loci were induced when the strain was grown on glucose, i.e., genes involved in metabolization of amylose (*BL0527 [malQ*]) and in transport of glucose (*BL1631* [*glcP*]) (Table 2). These findings tentatively support the involvement of AraQ and MalR1 in the control of the transcriptional adaptation elicited by the different carbon sources we studied and would thereby corroborate the role of these regulators in genome-wide gene expression regulation in *B. longum* NCC 2705.

Table 2: Differential expression values of genes that are predicted to belong to the AraQ (BL0275) and MalR1 (BL0142) regulons in *B. longum* NCC 2705, as presented in RegPrecise. Genes showing a significant differential expression are indicated in bold. Induction during galactose growth is depicted by positive log2FC, whereas negative log2FC values represent induction during glucose growth.

	LOCUS TAG	GENE NAME	LOG2FC	FDR Q-VALUE	FUNCTION
ARAQ REGULATED LOCI	**BL0141**	***malE***	**0.728**	**0.007**	**Maltose/maltodextrin ABC transp., substrate binding protein**
	BL0272	***araA***	**1.218**	**0.006**	**L-arabinose isomerase**
	BL0273	*araD*	ND	NA	L-ribulose-5-phosphate 4-epimerase
	BL0274	***araB***	**1.932**	**0.003**	**Ribulokinase**
	BL0275	***araQ***	**-1.106**	**0.037**	**Transc. regulator of central carb. metabolism, LacI family**
	BL0715	***tal***	**0.705**	**0.006**	**Transaldolase**
	BL0716	***tkt***	**0.526**	**0.011**	**Transketolase**
	BL0988	*pyk*	0.387	0.057	Pyruvate kinase
	BL0999	***glgB***	**0.697**	**0.002**	**GH13 glycosyl hydrolase**
	BL1000	-	1.388	0.091	Response regulator of two-component system
	BL1001	-	0.368	0.391	histidine kinase sensor of two-component system
	BL1022	*eno*	0.031	0.928	**Enolase**
	BL1308	*ldh*	-0.085	0.785	L-lactate dehydrogenase
	BL1359	***galM***	**2.006**	**0.002**	**Aldose 1-epimerase**
	BL1363	***gap***	**0.912**	**0.000**	**NAD-dependent glyceraldehyde-3-phosphate dehydrogenase**
	BL1531	**-**	**1.185**	**0.000**	**Regulator of polyketide synthase expression**
	BL1532	**-**	**1.581**	**0.000**	**hypothetical protein**
	BL1533	**-**	**-3.372**	**0.000**	**Biotin-protein ligase**
	BL1570	*malQ1*	0.168	0.590	4-alpha-glucanotransferase
MALR1 REGULATED LOCI	**BL0279**	***pgi***	**2.079**	**0.000**	**Glucose-6-phosphate isomerase**
	BL0597	*glgP*	-0.065	0.841	Glycogen phosphorylase
	BL0142	*malR1*	0.370	0.252	Transcr. regulator of maltose/maltodextrin utilization, LacI family
	BL0141	***malE***	**0.728**	**0.007**	**Maltose/maltodextrin ABC transp., substrate binding protein**
	BL1630	***pgm***	**1.006**	**0.001**	**Phosphoglucomutase (EC 5.4.2.2)**
	BL1631	***glcP***	**-0.930**	**0.001**	**D-Glucose-proton symporter**
	BL0527	***malQ***	**-1.834**	**0.000**	**4-alpha-glucanotransferase (amylomaltase)**
	BL0528	*malR2*	-0.349	0.664	Transcr. regulator of maltose/maltodextrin utilization, LacI family
	BL0529	*aglA*	-0.925	0.063	Alpha-glucosidase
	BL0143	*malF*	0.713	0.222	Maltose/maltodextrin ABC transporter
	BL0144	*malG*	0.699	0.263	Maltose/maltodextrin ABC transporter

Galactose grown *B. longum* NCC 2705 cells are more acid-stress resistant. It has previously been suggested that smaller bifidobacterial cells display an enhanced stress tolerance [233,234]. In addition, changes in acetate:lactate ratio [235], and expanded fermentative repertoires (i.e., catabolic flexibility) [236,237] have been correlated to an increased pH resistance. Our NCC 2705 transcriptome data indicated that several stress related genes were upregulated when the strain was grown on galactose. For example, the extracytoplasmic function (ECF) sigma factor (σ^E) (*BL1357*) and its putative anti-sigma factor (*BL1358*) were induced (Table S 1), and have previously been demonstrated to be part of the heat stress response of NCC 2705 [238] and the bile stress response of *B. longum* BBMN68 [239]. Notably, in several bacteria ECF-σ factors have also been shown to be important regulators of diverse stress responses [240,241], including the response to acid stress in *Streptococcus mutans* [242]. In addition, a set of genes predicted to be involved in the synthesis of rhamnose rich exopolysaccharides (EPS) were transcriptionally increased in galactose grown NCC 2705, including glucose-1-phosphate thymidylyltransferase (*BL0227*), dehydrorhamnose reductase (*BL0228*), EPS export genes (BL0207, BL0208), and extracellular glycosyltransferases (*BL0566, BL1674*) (Table S 1). Homologues of these genes were previously shown to be upregulated upon acid stress in *B. longum* BBMN68 and were proposed to support the maintenance of cell-integrity under this condition [243]. Taken together, these changes may predict improved acid-stress resistance in galactose grown NCC 2705 compared to their glucose grown counterparts. To assess this hypothesis, we performed a range of specific cell-stainings combined with flow-cytometry evaluations.

Initially, using propidium iodide (PI) stained cells, we observed that galactose grown cells intrinsically were significatively less permeable as compared to glucose grown cells (Figure 7 A). Subsequently, resistance of these *B. longum* NCC 2705 cultures to low pH stress was assessed by following the fraction of metabolically active cells over time during low pH exposure. To establish an appropriate pH stress condition, glucose grown cells harvested during the stationary phase were exposed to a pH ranging from 2.9 to 3.5 (using phosphate-buffered solutions), demonstrating that the rate of metabolic activity loss in the population of cells was proportional to the pH they were exposed to. This is illustrated by the observation that more than 85% of the cells failed to show any metabolic activity after 18 minutes of incubation at pH 2.9, while only 44%

had lost their metabolic activity upon exposure to pH of 3.5 for the same incubation time (Figure S 2). Thus, glucose or galactose grown *B. longum* NCC 2705 cells were harvested at the mid-logarithmic or stationary- phase of growth and exposed to a pH of 2.9 for a period of 18 min with regular assessment (every 3 minutes) of remaining metabolic activity in the population. A rapid loss of metabolic activity was observed in glucose grown cells, and this loss was faster in cells harvested from the mid-exponential ($k_{glu\text{-}expo}$=0.43) as compared to stationary ($k_{glu\text{-}stat}$=0.33) phase of growth (Figure 7 B and C). In contrast, loss of metabolic activity of mid-exponentially harvested galactose grown cells appeared to be delayed compared to their glucose grown counterpart, i.e., the decrease of metabolic activity only started after 6 minutes of pH 2.9 exposure. Moreover, galactose grown cells harvested during the stationary phase appeared more stress resistant since the loss of metabolic activity was even further delayed and a substantial proportion of the cells maintained their metabolic activity (~60 %) during the entire assay period. Similar to what was observed for glucose grown cells, metabolic activity loss in galactose grown cells declined at a faster rate in cells obtained from the exponential- ($k_{gal\text{-}expo}$=0.49) compared to the stationary- ($k_{glu\text{-}stat}$=0.33) (Figure 7 B and C). These measurements showed that, as expected, stationary phase harvested cells were more resistant compared to logarithmic phase harvested cells irrespective of the carbon source they were grown on. Moreover, these results show that irrespective of the growth phase of harvesting, glucose grown cells not only display a higher intrinsic permeability but also are more susceptible to pH stress (pH of 2.9) as compared to their galactose grown counterparts.

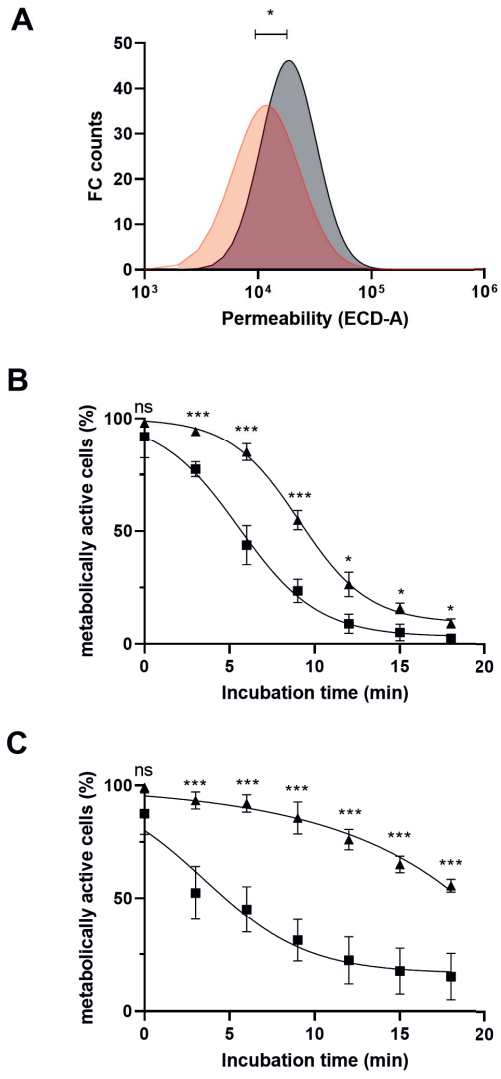


Figure 7: Cell permeability and acid sensitivity profile of *B. longum* NCC 2705 grown on glucose or galactose. Panel A shows the Propidium Iodide (PI) stained cells distribution of mid-exponential phase harvested glucose (grey) or galactose (red) grown *B. longum* NCC 2705. Acid sensitivity profile of *B. longum* NCC 2705 grown on glucose (squares) or galactose (triangles) when exposed to pH 2.9 is depicted in panels B and C. The loss of metabolically active cell populations (changing from CFDA+ to CFDA-) collected from exponentially (panel B) or the stationary (panel C) phase of growth is represented by the average and standard deviation of biological triplicates. Logistic growth modelling was applied to each curve of CFDA- population (r^2 >0.90) and inversed data were plotted. Stars represent the q-value (ns >0.05, *q <0.05; ***q<0.001) obtained upon multiple unpaired t-test for each time points.

Conclusions & discussion

Our results highlight important physiological differences between cells of *B. longum* NCC 2705 grown on glucose or galactose as sole carbon source. Compared to glucose, galactose grown NCC 2705 displayed a reduced growth rate and a smaller cell size. Gene transcription analysis revealed metabolic adaptations in glucose growing cells that may reflect the higher growth rate on this sugar, such as the upregulation of nucleotide biosynthesis and ATP generation. On the other hand, reduced growth rate on galactose was accompanied by an expanded catabolic flexibility phenotype, where a variety of transcriptional changes related to carbon source import and central energy metabolism supported our experimental findings. The two growth conditions were associated with an adaptation of the metabolic end-products, including an increased acetate:lactate ratio and the production of ethanol and formate by cells grown on galactose. Moreover, we demonstrated that galactose grown *B. longum* NCC 2705 were less permeable and displayed an enhanced acid-stress resistance as compared to glucose grown cells.

Catabolic flexibility, a bacterial physiological status under which a wider range of substrates can be utilized [231], is often induced in cells with a reduced growth rate [230]. Furthermore, growth rate reduction has been shown in different bacterial species to govern their resistance to different stresses [224-226,244,245]. A few studies using bifidobacteria have evidenced some of these features. For example, *B. longum* BBMN68 pre-exposed to mild acid stress exhibited an enhanced level of stress resistance, as well as increased expression of genes involved in carbohydrate metabolism and energy production [236]. Another study showed that acid-resistant derivatives of *B. longum* and *B. catenulatum* displayed increased catabolic flexibility [237]. Bile salt resistant derivatives of *B. bifidum* CECT 4549 and *B. infantis* CECT 4551 displayed an enhanced fructose-6-phoshoketolase (F6PPK) activity that was accompanied by an increased ratio of the metabolic end-products acetate:lactate [246]. Finally, acid resistance in *B. breve* BB8 was accompanied by reduced expression of genes involved in the production of lactic acid and increased expression of genes associated with ethanol production [235]. In this study, we conducted a head-to-head comparison of *B. longum* NCC 2705 grown on glucose or galactose. This allowed us to establish that galactose grown cells displayed a reduced growth rate accompanied by expansion of catabolic flexibility, modified central carbon metabolism (increased

ethanol and acetate:lactate ratio) and increased pH stress resistance. The transcriptome analysis we performed on these cells supports most of the observed phenotypic changes and points to the existence of a catabolite repression-like regulatory mechanism. We propose this mechanism to involve the two global gene expression regulators known in *B. longum* NCC 2705 (AraQ and MalR1 [181,182]), which control genes associated with the enhanced catabolic flexibility as well as the modified central carbon metabolism we observed in galactose grown cells. The transcriptome data imply that AraQ itself is produced at a lower level during growth on galactose as compared to glucose, suggesting that this decrease in AraQ expression may relieve the repression of several target genes, thus supporting its global role in carbon-source regulation in *B. longum* NCC 2705. In addition, the transcriptome revealed that galactose induced expression of the ECF σ^E (*BL1357*) and its postulated anti-sigma factor (*BL1358*) encoding genes, which we suggest to play a role in the elevated pH stress tolerance we observe. This hypothesis is in agreement with observations in other bacteria [247], like *E. coli* [224], *B. subtilis* [248] and *Streptococcus mutans* [242], where homologues of this alternative sigma factor were shown to control a large set of general stress response proteins and contribute to tolerance to a variety of stress conditions, including pH stress [242]. Therefore, we postulate that the induction of ECF σ^E in galactose grown NCC 2705 provides protection to a wider range of stress conditions, potentially including heat [238] and oxidative stress [59].

Galactose grown NCC 2705 produced elevated levels of ethanol, which contributes to NAD+ regeneration [45,51]. In other species, such as *L. lactis,* reduction of growth rate has been associated with a change in fermentation end products, including ethanol formation, and was additionally accompanied by an increased secondary metabolism activity [244]. Intriguingly, molecules that have been demonstrated to be important for the survival or host-interaction capacity of bifidobacteria, such as serpin [140] and sortase-dependant pili [129,135,136] have been shown to be regulated by different carbohydrates, warranting further research into the relationship between reduced growth rates and the production of bifidobacterial niche and effector molecules.

Overall, our results provide evidence for a prominent role of carbohydrate substrates in eliciting a broad spectrum of physiological changes in *B. longum* NCC 2705. The changes induced by growth on galactose are potentially very important for the adaptation of NCC 2705 to the gut environment, particularly in the early infancy, where

bifidobacteria are exposed to and catabolize different indigestible HMOs, many of which contain galactose moieties [169]. Moreover, glucose is likely to be much less abundant in the infant colon. The main source of glucose in human milk is in the form of lactose that is hydrolyzed by intestinal lactase, thereby releasing glucose which is rapidly absorbed in the small intestine of the infant [249]. Therefore, the galactose growth associated phenotype of NCC 2705 may arguably mimic the strain physiological status the infant gut, where catabolic flexibility and stress resistance enhancement are likely to increase the strain fitness. More generally, reduced growth rates (e.g., as induced by growth on galactose) may lead to the production of higher levels of secondary metabolites in the gut (i.e., niche factors and effector molecules).

Beyond their relevance to the bifidobacterial lifestyle in the gut, our results might have important implications for probiotic production and formulation. Although glucose is an attractive carbon source for industrial manufacturing due to its capacity to support fast growth, it may lead to the production of a bacterial biomass not optimally adapted to the intestinal milieu. Biomass produced using alternative carbohydrates as substrate (e.g., galactose) may display more appropriate attributes, preparing the cells for the challenging gastrointestinal conditions. As an alternative, the formulation of *Bifidobacterium* strains in combination with selected prebiotics (e.g., galacto-oligosaccharides or HMOs) that are rich in galactose moieties may increase the probiotic *in situ* adaptation. These considerations indicate that carbohydrate substrate selection for industrial production or the prebiotic added in the final synbiotic formulation, may modulate the physiology of the probiotic bacteria, which may in turn affect their *in situ* survival and efficacy.

Materials & methods

***B. longum* NCC 2705 cultivation.** *B. longum* NCC 2705 was obtained from the Nestle Culture Collection (NCC) (Nestlé Research, Lausanne, Switzerland). The strain was grown in a modified de Man, de Rogosa & Sharpe medium lacking carbohydrates(MRS-C) (10 g/L bacto proteose peptone, 3, 5 g/L bacto yeast extract, 1 g/L Tween 80 [Chemie Brunschwig, Basel, Switzerland]; 2 g/L di-ammonium hydrogen citrate, 5 g/L sodium acetate, 0.1 g/L magnesium sulphate, 0.05 g/L manganese sulfate, 2 g/L di-sodium phosphate, 0.5 g/L cysteine [Sigma-Aldrich Chemie GmbH, Buchs, Switzerland]). MRS-C was supplemented with 1% galactose or 1% glucose (Sigma-Aldrich Chemie GmbH) using filter sterilized (0.22 µm) stock solutions prepared at 100 g/L in water. Overnight cultures were inoculated at 2 % in fresh medium and growth was performed at 37°C in 2 liter DasGip bioreactors (Eppendorf AG, Hamburg, Germany), using 500 ml working volume. Anaerobiosis was achieved by CO_2 flux in the headspace, and agitation was maintained at 600 rpm. pH was monitored throughout the fermentation and samples were taken at regular interval to measure Optical Density (OD measured at 600 nm) and enumerate Colony Forming Units (CFU) using standard MRS-agar supplemented with 0.05 % cysteine (48h anaerobic incubation at 37°C). In addition, at the same timepoints cells were collected by centrifugation (3000 g, 10 min at 4°C) and resuspended in sterile PBS supplemented with 15% glycerol (Sigma-Aldrich Chemie GmbH). Aliquots were stored at -80°C until further analysis. Every fermentation condition was run in triplicate.

Growth data analysis. Raw data obtained from the above fermentations were analysed as follows. Maximal acidification slope was obtained by modeling inversed pH values with a logistic growth model in GraphPad Prism (v8.1.1, GraphPad Software Inc., San Diego, USA). The same modelling strategy was applied to the OD 600 nm values to obtain the maximal growth rate in each fermentation. All CFU/ml and OD values obtained before the decrease in viability (in each bioreactor) were used to calculate the average CFU:OD ratios. For each condition, OD values were multiplied by the corresponding CFU:OD ratio to obtain theoretical CFU/ml. A logistic growth model was then applied on the theoretical CFU/ml values, to finally estimate the cell doubling time for each condition.

Microscopic observations. Samples collected at the different stages of growth (mid-exponential, early and late stationary phases) were fixed using a 10% solution of Nigrosin (Sigma-Aldrich Chemie GmbH) and analysed by confocal microscopy (Olympus Life Science, Hamburg, Germany). For each growth phase and growth condition, the images were analysed using the cellSens Entry software (v1.16) (Olympus Life Science) to determine the cell size of a minimum of 50 individual cells. All data points were then merged (per growth condition) to calculate an averaged cell length and the respective standard error of the mean (SEM) in GraphPad Prism (v8.1.1, GraphPad Software Inc.).

RNA isolation and sequencing. Cells were harvested from 2 ml culture samples collected during mid-exponential growth (OD_{600} ~ 0.6-0.8) by centrifugation (3000 g, 3 min at 4°C). RNA was isolated from each biological triplicate using the miRNeasy kit (QIAGEN AG, Hombrechtikon, Switzerland). RNA concentration was measured using the Quant-it Ribogreen RNA kit (Thermo Fisher Scientific, Waltham, USA) and RNA quality was evaluated using the Agilent 2100 Bioanalyser system with the Agilent RNA 6000 nano kit (Agilent Technologies, Santa Clara, USA). Library preparations were performed with the Prokaryotic AnyDeplete kit, which provides all reagents necessary to produce ready to sequence RNA libraries (Protocol M01502 v1, NuGEN Technologies Inc., Redwood City, USA). The Prokarotic AnyDeplete workflow is a follows: 250 ng of total RNA was used as input to produce cDNA, followed by an enzymatic fragmentation to a target size of around 200pb. End-repair of the cDNA was performed, and specific adaptors provided with the kit (8 bases index allowing sample multiplexing for Illumina sequencing) were added by ligation. A strand selection step (to retain RNA strand information) was followed by a ribodepletion step aiming at eliminating unwanted rRNA transcript sequences. Finally, 18 PCR cycles were used for the final amplification of the libraries using specific primers provided in the AnyDeplete kit. Obtained DNA fragments were then purified with a 1:1 ratio of AMPURE XP beads (Beckman Coulter Life Sciences, Indianapolis, USA). Libraries were checked on a Labchip GX touch with DNA High Sensitivity Reagent Kit (PerlinElmer Inc., Waltham, USA). All samples were finally sequenced on an Illumina HiSeq 2500 with Rapid V2 chemistry (Illumina Inc., San Siego, USA), applying 200 cycles (paired-end reads PE100). Flowcell was loaded with 10 pM of DNA and with

2% PhiX control v3 library (internal control enabling quality check of the sequencing run; Illumina Inc.).

Transcriptome analysis. A dedicated RNAseq analysis pipeline for prokaryotes [250] was used for the analysis of the obtained *B. longum* NCC 2705 transcriptional data. Following mapping with HISAT2 (version 2.1.0)[251], reads were softclipped and mapping quality set to 0 for unmapped reads (picard-tools version 1.119, CleanSam; https://gatk.broadinstitute.org/hc/en-us/articles/360036885571-CleanSam-Picard-). The resulting bam files were then merged using samtools [252] (version 1.4) prior to read counting using FADU [250] (version 1.7). Counts obtained were first filtered to keep only genes having at least 5 read-counts in at least 2 samples. Normalization was performed using RLE (Relative Log Expression, using the DEseq2 R package [253]) to compensate for composition bias between the different libraries. Differential expression analysis was performed using edgeR [254]. Differentially expressed genes (DEG) were filtered based on their False Discovery Rate corrected p-values (FDR<0.05) and used for Gene Ontology (GO) enrichment, which was performed with the BINGO Cytoscape plugin [255], using the B. longum NCC 2705 reference proteome available at Uniprot as basis for annotation [256]. A central carbon metabolism pathway map was reconstructed in Cytoscape based on different KEGG maps (00052, galactose metabolism; 00030, pentose phosphate; 00010, glycolysis; 00620, pyruvate metabolism). Similarly, a map of the different sugar transporter systems present in the strain was reconstituted starting from Parche et al [19]. Finally, the log2FC values obtained previously were mapped on the list of AraQ and MalR1 regulons retrieved from the RegPrecise database [180], and used to color the constructed maps in Cytoscape.

Acetate, lactate, ethanol and formate quantifications. The (relative) levels of formate, acetate, lactate and ethanol in spent culture supernatants were determined by NMR spectroscopy. To this end, *B. longum* NCC 2705 was cultured to mid-exponential phase on glucose and galactose, in thighly closed tubes. Cells were removed by centrifugation (3300 g, 2 min at 4°C) and the culture supernatant was filtered through a 0.22 µm polyether sulfone filter and stored at -20 °C until further analysis. For NMR spectroscopy samples were diluted 1:1 in a 10 % D_2O/90 % H_2O solution. 1D Nuclear Overhauser Effect Spectroscopy (NOESY) spectra were measured on a 600 MHz NMR spectroscopy (Bruker BioSpin, Billerica, USA) equipped

with a 5 mm cryo-probe at a temperature of 300 K. The ^{1}H-^{1}H mixing time was 10 ms. A saturation of 50 Hz was applied on water during relaxation delay and mixing. Quantum Mechanics-based 1H full Spin Analysis with the Cosmic Truth software (NMR Solutions Ltd., Kuopio, Finland) was employed for data analysis and peak assignments (for details see [257,258]). Peak intensities of the metabolic end-products were quantified in comparison to spiked-in standard of ^{13}C labelled acetate.

Flow cytometry analysis. Flow cytometry analyses were performed with a Beckman Coulter Cytoflex S (Beckman Coulter, Brea, US) equipped with four lasers. Bacterial cell staining was performed using 1.5 µM PI and 5 µM CFDA with an incubation of 10 minutes at room temperature before analysis. Forward scatter acquisition threshold was set at 1000 in order to capture all cells, including the small galactose grown ones. Cell size was estimated using the Forward Scatter signal (FSC). Cell permeability was evaluated through propidium Iodide (PI) (Thermo Fisher Scientific AG) staining and measured using the ECD channel (excitation at 561nm, emission at 610/20 nm). Metabolic activity was evaluated using the cell-permeant esterase substrate 5-carboxyfluorescein diacetate acetocymethyl ester (CFDA) (Thermo Fisher Scientific AG) measured using the FITC detector (excitation at 488 nm, emission at 525/40 nm). Acquisitions were processed with the Beckman-Coulter CytExpert Acquisition software version 2.3, followed by further analysis with the FCSExpress software version 7.01 (De Novo Software, Pasadena, USA). Statistical analysis of population distribution was performed in GraphPad Prism (v8.1.1, GraphPad Software Inc.), using a welch-t test.

Translationally blocked acidification assay. The assay used was adapted from Nugroho et al [232], and aimed to determine the acidification capacity of translationally blocked glucose or galactose grown B. longum NCC 2705 using a set of different carbohydrate substrates. The assay was performed in microwell plates under anaerobic conditions. Protein translation was blocked using erythromycin and once the bacteria switched to the new carbohydrates, acidification was followed over the incubation period using carboxyfluorescein. The assay medium was adapted from MRS-C by removing yeast extract, to avoid contamination steming from the residual sugars contained in this ingredient.

For this assay, *B. longum* NCC 2705 was grown to the mid-exponential phase on MRS-C supplemented with either glucose or galactose as a sole carbon source. Five ml of cultures were harvested by centrifugation (2 min at 3300 g), washed with 1 volume of sterile PBS and resuspended at an OD_{600} of 0.6 (final OD in the well 0.3) in 2X concentrated assay medium without sugar (for 1X: 5 g/L bacto proteose peptone, 1 g/L Tween 80 [Chemie Brunschwig]; 2 g/L di-ammonium hydrogen citrate, 5 g/L sodium acetate, 0.1 g/L magnesium sulfate, 0.05 g/L manganese sulfate, 2 g/L di-sodium phosphate, 0.5 g/L cysteine, 5 µg/ml erythromycin, 10 µM carboxyfluorescein [Sigma-Aldrich Chemie GmbH]). 100 µl of cell suspension was added to the different wells, which were all prefilled with 100 µl of the different 2X filtered sterile carbohydrate solutions (final sugar concentration in the well of 5 g/L; glucose, galactose, fructose, raffinose, maltose, lactose, sucrose, arabinose [Sigma-Aldrich Chemie GmbH], NutraFlora scFOS [Ingredion Korea Inc, Gyunggi-do, Korea], Vivinal GOS [FrieslandCampina DOMO Amersfoort, The Netherlands], Lacto-N-tetraose LNT [Glycom A/S, Lyngby, Denmark]). All solutions (2X concentrated assay medium, 2x carbohydrate solutions) were prewarmed at 37°C in a water bath before starting the assay. Finally, the 96 microwell plate containing the translationally blocked cells mixed with the different sugars (200 µl in each well) was sealed in a plastic bag containing an AGELESS ZPT mini anaerobic sachet (Mitsubishi Gas Chemical Gmbh, Dusseldorf, Germany) and immediately transferred in a Varioskan spectrophotometer (Thermo Fisher Scientific AG), maintained at a temperature of 37°C. Fluorescence was measured every 10 min (excitation at 485 nm, emission at 535 nm) over a period of 200 min. A logarithmic curve was then fitted to inversed individual acidification curves in excel, and acidification rates were extracted by calculating of the slope of the curve, followed by data transfer to GraphPad Prism (v8.1.1, GraphPad Software Inc.) where they were subjected to an unpaired multiple t-test comparison.

Loss of metabolic activity upon exposure to low pH. Frozen mid-exponential or stationary phase harvested cells of B. longum NCC 2705, stored at -80°C in a PBS suspension containing 15% glycerol, were used as starting material. Initial trials aimed to select an optimal pH under which a significant fraction of the cells in the suspension lost their metabolic activity during a stress exposure period of 18 minutes. To assess the appropriate incubation pH to achieve the desired assay conditions, glucose grown stationary phase harvested cells were thawed on ice, washed in one volume of sterile

PBS (centrifugation 2 min, 3300 g), and resuspended to an OD of 1.0 in PBS adjusted at different pH values (pH 2.9, 3.0, 3.2, 3.5) using 1 M hydrogen chloride (HCl). The obtained solution was distributed in 100 µl aliquots and incubated at 37°C for a period of 18 minutes. An aliquot was collected every 3 min, and 900 µl of PBS (pH 6.7) containing 1.5 µM PI and 5 µM CFDA were added to stop the pH induced loss of metabolic activity. After an incubation of 10 minutes at room temperature permeability and metabolic activity of at least 5000 individual cells was measured by flow cytometry (see above) at each timepoint and for each condition. The increase in non-metabolically active cell population fractions (CFDA-) over time was then modelled in GraphPad Prism (v8.1.1, GraphPad Software Inc.) using a logistic growth curve, as recommended by de Besten et al. [197]. Loss of metabolic activity in glucose and galactose grown cultures (mid-exponential and stationary) upon exposure to a pH of 2.9 was then pursued using a similar protocol. The assay was performed using stored culture samples from duplicate fermentations, and measurements for each sample were repeated to obtain technical duplicates. Additionally, frozen cells were compared to freshly obtained cultures to ensure that the freezing process did not impact the resistance phenotypes.

Supplementary data

Table S 1: List of differentially expressed *B. longum* NCC 2705 genes (FDR<0.05) upon growth on galactose (positive log2FC) or glucose (negative log2FC values) as sole carbon source. Gene annotation was retrieved from www.uniprot.org. FDR represents the p-value after false discovery rate application.

Gene ID	Gene name	Uniprot annotation	log2FC	FDR
BL1398		Uncharacterized protein	-3,51	6,37E-04
BL0745		Uncharacterized protein	-3,49	4,54E-04
BL1533		BPL/LPL catalytic domain-containing protein	-3,37	2,65E-38
BL0476		Narrowly conserved hypothetical membrane protein	-3,36	5,70E-61
BL1163		Probable solute binding protein of ABC transporter system for sugars	-3,24	7,29E-27
BL1133		Protein with weak similarity to components of ABC transporter	-3,15	5,30E-55
BL1742		Ppx-GppA domain-containing protein	-2,87	8,08E-18
BL0573		Uncharacterized protein	-2,68	4,84E-20
BL1802		Uncharacterized protein	-2,67	2,14E-14
BL1692		ATP binding protein of ABC transporter for pentoses	-2,56	1,62E-40
BL0491		Uncharacterized protein	-2,43	1,86E-11
BL1303		Hypothetical transmembrane protein possibly involved in transport	-2,40	9,62E-18
BL0087	dapF	Diaminopimelate epimerase (EC 5.1.1.7)	-2,39	6,57E-29
BL0799		Uncharacterized protein	-2,39	3,42E-10
BL0765		N-acetyltransferase domain-containing protein	-2,37	3,39E-07
BL1161		Solute binding protein of ABC transporter systemy for peptides	-2,31	7,06E-09
BL1256		Uncharacterized protein	-2,29	7,59E-20
BL1422	trpD	Anthranilate phosphoribosyltransferase (EC 2.4.2.18)	-2,28	6,03E-10
BL0298	stbB vapC	Ribonuclease VapC (RNase VapC) (EC 3.1.-.-)	-2,19	8,15E-06
BL0896		VOC domain-containing protein	-2,19	1,82E-07
BL0677		Uncharacterized protein	-2,18	9,00E-14
BL0302		Uncharacterized protein	-2,10	3,34E-12
BL1067	pheS	Phenylalanine--tRNA ligase alpha subunit (EC 6.1.1.20)	-2,06	2,11E-03
BL0897	pncA	Probable pyrazinamidase/nicotinamidase	-2,05	5,04E-11
BL0727	ruvC	Crossover junction endodeoxyribonuclease RuvC (EC 3.1.22.4)	-2,04	2,07E-02
BL1530		Probable lclR-type transcriptional regulator	-2,02	1,28E-23
BL0161		Uncharacterized protein	-2,00	1,30E-27
BL0195		Glutamine amidotransferase type-2 domain-containing protein	-1,99	1,75E-03
BL1150	yfnA	Probable amino acid transporter	-1,99	9,32E-14
BL0323	pcp	Pyroglutamyl-peptidase I (EC 3.4.19.3)	-1,98	7,98E-03
BL0263		Solute binding protein of ABC transporter system	-1,98	2,22E-04
BL0042		CRM domain-containing protein	-1,97	5,00E-14
BL1686		Uncharacterized protein	-1,92	1,51E-21
BL0489		Uncharacterized protein	-1,92	1,46E-07
BL0878	aroC	Chorismate synthase (CS) (EC 4.2.3.5)	-1,91	1,58E-04
BL0925		Uncharacterized protein	-1,90	1,98E-10
BL1497		Peptidase_M23 domain-containing protein	-1,86	1,19E-15
BL0102		Possible inosine-uridine preferring nucleoside hydrolase	-1,86	9,65E-27
BL1613		LacI-type transcriptional regulator	-1,85	3,70E-09
BL0226		Uncharacterized protein	-1,84	2,25E-03
BL0527	malQ	4-alpha-glucanotransferase (EC 2.4.1.25)	-1,83	9,56E-07
BL0092		DNA helicase (EC 3.6.4.12)	-1,83	4,60E-07
BL1644	galE1	UDP-glucose 4-epimerase (EC 5.1.3.2)	-1,82	1,17E-11
BL0801		Abhydrolase_2 domain-containing protein	-1,79	1,40E-08
BL0114	thiE/thiC	Thiamine biosynthesis bifunctional protein ThiEC (EC 2.5.1.3)	-1,79	6,06E-08
BL0707	pgk	Phosphoglycerate kinase (EC 2.7.2.3)	-1,78	2,93E-04
BL0996		Possible permease of ABC transporter system	-1,78	1,46E-05
BL1536	pccB	Propionyl-CoA carboxylase beta chain	-1,75	1,76E-02
BL0921		Uncharacterized protein	-1,74	6,25E-12
BL0039		Uncharacterized protein	-1,72	8,70E-03
BL0390	mntP	Putative manganese efflux pump MntP	-1,71	6,12E-04
BL1094		UPF0324 membrane protein BL1094	-1,71	4,01E-13
BL0171		Uncharacterized protein	-1,69	2,84E-05
BL1276		Nucleoside triphosphate pyrophosphatase (EC 3.6.1.9)	-1,68	3,57E-04
BL0846		LacI-type transcriptional regulator	-1,66	2,06E-05
BL0719	dnaJ	Chaperone protein DnaJ	-1,66	4,41E-14
BL1488		Uncharacterized protein	-1,66	1,45E-05
BL0809		Uncharacterized protein	-1,64	9,22E-17
BL1427		Uncharacterized protein	-1,63	1,43E-02
BL0720		Uncharacterized protein	-1,63	4,61E-22
BL1719		Probable sugar O-acetyltransferase (thiogalactoside acetyltransferase)	-1,59	4,29E-07
BL1740		Probable oxidoreductase	-1,59	2,22E-13
BL1046		Uncharacterized protein	-1,58	1,22E-07
BL1727		Uncharacterized protein	-1,57	4,82E-02
BL0778	hutM	Histidine permease	-1,55	1,58E-07

BL1541		Uncharacterized protein	-1,53	2,39E-08
BL1278		Uncharacterized protein	-1,53	2,78E-07
BL1326	dtd	D-aminoacyl-tRNA deacylase (DTD) (EC 3.1.1.96)	-1,53	8,28E-03
BL1424		PlsC domain-containing protein	-1,51	9,47E-03
BL1078		Possible integral membrane permease protein	-1,47	1,11E-12
BL0614		Thioredoxin reductase-like protein	-1,46	3,32E-12
BL0574		Uncharacterized protein	-1,45	1,74E-03
BL0164		FAD:protein FMN transferase (EC 2.7.1.180)	-1,43	1,16E-15
BL0737	rluB	Pseudouridine synthase (EC 5.4.99.-)	-1,43	1,14E-02
BL0328		Uncharacterized protein	-1,37	5,17E-05
BL0898		Uncharacterized protein	-1,36	5,34E-03
BL0253		Uncharacterized protein	-1,33	3,48E-05
BL1399		PINc domain-containing protein	-1,32	9,94E-05
BL0089		POLIIIAc domain-containing protein	-1,31	6,71E-14
BL1189		Uncharacterized protein	-1,30	2,28E-02
BL1295	hisD	Histidinol dehydrogenase (HDH) (EC 1.1.1.23)	-1,30	3,58E-02
BL1355		Acetyltransf_6 domain-containing protein	-1,29	9,68E-03
BL0913		HTH cro/C1-type domain-containing protein	-1,27	3,65E-02
BL0403	gatA	Glutamyl-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.7)	-1,26	4,13E-04
BL1311		LysM domain-containing protein	-1,25	2,95E-03
BL1052		Uncharacterized protein	-1,24	1,66E-13
BL1565	dinB	DNA polymerase IV (Pol IV) (EC 2.7.7.7)	-1,23	4,68E-10
BL0349		CYTH domain-containing protein	-1,23	9,69E-08
BL1361	ispH	4-hydroxy-3-methylbut-2-enyl diphosphate reductase (EC 1.17.7.4)	-1,23	3,37E-09
BL1553		SAF domain-containing protein	-1,22	4,24E-12
BL0071	dnaQ	Possible exonuclease	-1,20	9,74E-08
BL0063		Uncharacterized protein	-1,20	1,13E-07
BL1356	murE	UDP-N-acetylmuramyl-tripeptide synthetase (EC 6.3.2.-)	-1,20	1,66E-02
BL0013	proP	Proline/betaine transporter	-1,20	2,44E-02
BL0084		PHB domain-containing protein	-1,20	1,70E-10
BL0443		Probable 6-phosphogluconolactonase	-1,19	2,16E-09
BL1268		Probable aminotransferase	-1,19	4,71E-11
BL0288		Uncharacterized protein	-1,19	4,92E-04
BL0072		Tr-type G domain-containing protein	-1,19	7,03E-11
BL1785	hsdR	Type I restriction enzyme R Protein (EC 3.1.21.3)	-1,18	1,05E-06
BL0937		Uncharacterized protein	-1,17	5,87E-07
BL0439	pepDB	Dipeptidase (EC 3.4.-.-)	-1,16	2,30E-05
BL1137		Protein similar to hex regulon repressor	-1,16	3,83E-03
BL0639		Uncharacterized protein	-1,15	2,06E-07
BL1548		Uncharacterized protein	-1,15	2,60E-09
BL1186	def1	Peptide deformylase 1 (PDF 1) (EC 3.5.1.88)	-1,13	9,25E-03
BL0322		Uncharacterized protein	-1,13	2,71E-04
BL1354	smc	Chromosome partition protein Smc	-1,13	5,09E-08
BL1003		Pribosyltran domain-containing protein	-1,13	9,18E-08
BL1110		Uncharacterized protein	-1,12	3,46E-04
BL1650		Uncharacterized protein	-1,12	9,70E-05
BL1176		ATP binding protein of ABC transporter	-1,12	4,62E-02
BL0290		Possible reductase	-1,11	1,60E-08
BL0275		LacI-type transcriptional regulator	-1,11	3,68E-02
BL1537	fas	Fas	-1,10	3,93E-06
BL1048		Hemolysin-like protein with S4 domain found in bacteria and plants	-1,10	9,77E-03
BL1714		Solute binding protein of ABC transp. for branched-chain amino acids	-1,10	3,05E-03
BL0281		Uncharacterized protein	-1,10	4,90E-06
BL1344	nagA	N-acetylglucosamine-6-phosphate deacetylase	-1,09	2,07E-02
BL0699		Uncharacterized protein	-1,09	2,77E-02
BL1578	rplC	50S ribosomal protein L3	-1,09	5,63E-05
BL1199		Large protein with C-terminal fibronectin type III domain	-1,08	1,26E-07
BL0362	atpE	ATP synthase subunit c (ATP synthase F(0) sector subunit c)	-1,07	5,19E-05
BL1214	baiC	Probable NADH-dependent flavin oxidoreductase YqjM	-1,05	9,74E-08
BL0841		Uncharacterized protein	-1,04	1,28E-05
BL1267	murA	UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7)	-1,02	9,56E-07
BL0890	pcrA	DNA helicase (EC 3.6.4.12)	-1,02	2,43E-02
BL1197		PDDEXK_1 domain-containing protein	-1,02	1,54E-06
BL1031		Possible B-hexosaminidase	-1,01	2,15E-08
BL0202		Uncharacterized protein	-1,00	1,85E-03
BL0009		Uncharacterized protein	-1,00	4,93E-08
BL1633		Transcription antiterminator similar to LicT	-0,99	2,71E-05
BL1100	rpsL	30S ribosomal protein S12	-0,97	8,20E-03
BL0361	atpF	ATP synthase subunit b (ATP synthase F(0) sector subunit b)	-0,97	5,89E-06
BL1352		Uncharacterized protein	-0,97	3,61E-07
BL1124	aldH	Aldehyde dehydrogenase	-0,97	3,48E-03
BL0162		Probable permease protein of ABC transporter system	-0,96	1,79E-02
BL1423		Uncharacterized protein	-0,96	2,02E-02
BL1672		Possible cyclopropane-fatty-acyl-phospholipid synthase	-0,96	5,99E-05
BL0433	glnD	Protein-pII uridylyltransferase	-0,94	8,15E-06

BL1631	glcP	D-Glucose-proton symporter	-0,93	1,42E-03
BL0329	recG	ATP-dependent DNA helicase	-0,92	2,95E-02
BL1657		Phosphate-specific transport system accessory protein PhoU	-0,92	9,42E-03
BL1014		Uncharacterized protein	-0,90	1,36E-06
BL1547		Uncharacterized protein	-0,89	3,81E-03
BL0432		Uncharacterized protein	-0,89	1,33E-05
BL1406		Possible integral membrane protein with duf6	-0,89	3,39E-07
BL0607		Uncharacterized protein	-0,88	4,12E-02
BL0722		Uncharacterized protein	-0,88	1,79E-05
BL1227		G5 domain-containing protein	-0,86	1,94E-02
BL0419		Pyridoxal phosphate homeostasis protein (PLP homeostasis protein)	-0,86	2,06E-04
BL1066	pheT	Phenylalanine--tRNA ligase beta subunit (EC 6.1.1.20)	-0,85	1,15E-02
BL1336		Probable LacI-type transcriptional regulator	-0,85	3,99E-03
BL1379		TypA/BipA	-0,85	5,41E-03
BL0480		Uncharacterized protein	-0,84	3,09E-06
BL0151		Uncharacterized protein	-0,84	4,45E-05
BL0606		Uncharacterized protein	-0,83	2,77E-02
BL0930		Response regulator of two-component system	-0,83	1,40E-04
BL0358	atpG	ATP synthase gamma chain (ATP synthase F1 sector gamma subunit)	-0,83	2,50E-03
BL0046		TM1586_NiRdase domain-containing protein	-0,82	2,84E-05
BL1341		Possible sugar kinase	-0,80	6,84E-03
BL0721	uppP bacA	Undecaprenyl-diphosphatase (EC 3.6.1.27)	-0,80	2,07E-02
BL0011	codA	Cytosine deaminase	-0,80	3,03E-05
BL0883		Uncharacterized protein	-0,79	2,48E-06
BL1647		PspC domain-containing protein	-0,78	3,76E-03
BL1312	nrdR	Transcriptional repressor NrdR	-0,77	1,41E-02
BL1366	infC	Translation initiation factor If-3	-0,76	1,76E-02
BL0186		Possible esterase	-0,76	1,03E-02
BL1515		Probable membrane transporter protein	-0,76	1,17E-04
BL1528		Sir2-type regulatory protein	-0,75	2,04E-02
BL1743		Endo/exonuclease/phosphatase domain-containing protein	-0,75	5,29E-05
BL0919		Possible efflux transporter protein	-0,74	6,07E-03
BL0634	gyrA1 gyrA	DNA gyrase subunit A (EC 5.6.2.2)	-0,74	5,20E-03
BL0170	metG	Methionine--tRNA ligase (EC 6.1.1.10)	-0,74	2,80E-02
BL1527	tdcB	Catabolic threonine dehydratase	-0,74	3,27E-02
BL1291	rplA	50S ribosomal protein L1	-0,73	2,43E-02
BL1332		Probable sugar permease of ABC transporter system	-0,73	3,22E-02
BL0062		Protein similar to YcbI of B. subtilis	-0,71	3,37E-03
BL1257		Uncharacterized protein	-0,70	2,99E-03
BL0448		Possible TetR-type transcriptional regulator	-0,69	6,11E-04
BL0557		Uncharacterized protein	-0,69	4,63E-03
BL0050		Uncharacterized protein	-0,68	8,45E-05
BL0498		Possible sortase-like protein	-0,68	2,77E-02
BL0015		5'-nucleotidase family protein	-0,68	1,13E-02
BL0713		Uncharacterized protein	-0,67	2,32E-04
BL1410		Fido domain-containing protein	-0,67	2,53E-04
BL0982	glgX	Probable glycogen operon protein GlgX	-0,66	1,92E-02
BL0270		Uncharacterized protein	-0,65	1,02E-02
BL1594	rpsH	30S ribosomal protein S8	-0,65	5,14E-03
BL1140		Possible transport protein	-0,63	2,19E-04
BL0359	atpA	ATP synthase subunit alpha (EC 7.1.2.2)	-0,63	4,04E-03
BL1365	speE	Probable polyamine aminopropyl transferase (EC 2.5.1.16)	-0,63	7,38E-03
BL0073		Possible phospholipase/carboxylesterase	-0,62	1,82E-03
BL1006		Uncharacterized protein	-0,62	8,79E-03
BL0532	shiA	Transmembrane transport protein possibly for shikimate	-0,60	3,27E-02
BL1571	rplM	50S ribosomal protein L13	-0,59	6,06E-03
BL1725		AMMECR1 domain-containing protein	-0,59	3,73E-02
BL0914		Uncharacterized protein	-0,58	4,61E-03
BL1099	rpsG	30S ribosomal protein S7	-0,58	1,05E-02
BL0779	sipY	Signal peptidase I (EC 3.4.21.89)	-0,58	1,06E-03
BL0387		Uncharacterized protein	-0,58	4,19E-03
BL1432	lhr	ATP-dependent helicase II	-0,57	2,93E-02
BL1509		DJ-1_Pfpl domain-containing protein	-0,53	1,02E-02
BL0490		Metallophos domain-containing protein	-0,53	2,66E-02
BL1668	alzD	Branched-chain amino acid permease	-0,52	1,46E-02
BL0759		MFS domain-containing protein	-0,51	1,02E-02
BL0714		Branched-chain amino acid transport system carrier protein	-0,50	6,76E-03
BL0899		Uncharacterized protein	-0,49	1,84E-02
BL1386	dppA2	DppA2	-0,49	2,66E-02
BL1125		Sulfatase domain-containing protein	-0,49	3,27E-02
BL1425	pknA2	Probable serine/threonine-protein kinase pknA2 (EC 2.7.11.1)	-0,47	2,79E-02
BL0674		Possible cell surface protein with Gram-positive anchor domain	-0,45	3,42E-02
BL0266	fadD1	Probable long-chain-fatty acid CoA ligase	0,43	3,88E-02
BL1417		Uncharacterized protein	0,49	5,35E-03
BL1595	rplF	50S ribosomal protein L6	0,49	3,81E-02

BL0970	aroA	3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19)	0,52	2,55E-02
BL0716	tkt	Transketolase	0,53	1,08E-02
BL0213		Uncharacterized protein	0,53	1,74E-02
BL1362		Cys-tRNA(Pro)/Cys-tRNA(Cys) deacylase (EC 4.2.-.-)	0,53	7,06E-03
BL1543	xynD	Endo-1,4-beta-xylanase D	0,54	2,62E-02
BL1203		Uncharacterized protein	0,55	1,68E-03
BL1290	rplK	50S ribosomal protein L11	0,57	7,80E-03
BL0794	pyrB	Aspartate carbamoyltransferase (EC 2.1.3.2)	0,59	2,89E-02
BL1803		Uncharacterized protein	0,60	1,74E-02
BL1503	rpsB	30S ribosomal protein S2	0,60	7,95E-03
BL0874	pyrG	CTP synthase (EC 6.3.4.2) (Cytidine 5'-triphosphate synthase)	0,60	9,24E-03
BL0893		HTH cro/C1-type domain-containing protein	0,61	1,79E-02
BL1632	ptsG	PtsG	0,61	4,63E-03
BL1123	purD	Phosphoribosylamine--glycine ligase (EC 6.3.4.13)	0,61	3,81E-02
BL0199		Uncharacterized protein	0,62	3,56E-02
BL0855		Uncharacterized protein	0,62	3,34E-03
BL0821	bcp	Possible thioredoxin-dependent thiol peroxidase	0,62	2,16E-03
BL0739	ugpA	Probable UTP-glucose-1-phosphate uridylyltransferase	0,63	1,76E-02
BL0555	degP	Possible DO serine protease	0,65	6,64E-04
BL0400a		Uncharacterized protein	0,65	1,29E-04
BL0691		Uncharacterized protein	0,66	3,47E-02
BL1688		ATP binding protein of ABC transporter	0,68	1,79E-02
BL0289		MerR-type transcriptional regulator	0,68	2,41E-04
BL1076	glnA1	Glutamine synthetase (EC 6.3.1.2)	0,68	1,38E-02
BL0999	glgB	1,4-alpha-glucan branching enzyme GlgB (EC 2.4.1.18)	0,70	2,22E-03
BL0940	safC	Probable methyltransferase	0,70	2,56E-02
BL0715	tal	Transaldolase (EC 2.2.1.2)	0,70	6,07E-03
BL0683	pacL1	Cation-transporting ATPase PacL	0,72	1,74E-02
BL1322	fstW	Probable FtsW-like protein	0,72	1,66E-03
BL1659		Uncharacterized protein	0,72	8,61E-05
BL0141		Possible solute binding protein of ABC transporter	0,73	7,07E-03
BL0673	msiK	ATP binding protein of ABC transporter for sugars	0,73	3,19E-03
BL0005		Response regulator of two-component system	0,74	4,36E-02
BL1584	rpsC	30S ribosomal protein S3	0,74	4,08E-02
BL1396	ctpE	Probable cation-transporting ATPase	0,75	1,20E-03
BL1716		Permease of ABC transporter for branched-chain amino acids	0,76	2,96E-02
BL0450		ATP binding protein of ABC transporter	0,77	2,64E-02
BL1664a		Uncharacterized protein	0,77	1,59E-05
BL1251	gluQ	Glutamyl-Q tRNA(Asp) synthetase (Glu-Q-RSs) (EC 6.1.1.-)	0,78	5,99E-03
BL0372		HNHc domain-containing protein	0,79	1,87E-04
BL0168		Uncharacterized protein	0,79	2,04E-02
BL1691	glkA	Glucokinase	0,80	8,69E-05
BL0206		Uncharacterized protein	0,80	5,05E-06
BL0214		Uncharacterized protein	0,81	4,00E-05
BL1664		Widely conserved protein in universal stress protein family	0,82	1,33E-03
BL0167		Uncharacterized protein	0,82	3,73E-04
BL1220		DUF2183 domain-containing protein	0,83	8,31E-04
BL0575		EamA domain-containing protein	0,83	1,63E-05
BL1139		Uncharacterized protein	0,84	7,86E-03
BL1198		Probable serine/threonine-protein kinase	0,84	4,56E-02
BL0407		Uncharacterized protein	0,84	1,88E-02
BL1460		Uncharacterized protein	0,84	2,60E-02
BL1405		Uncharacterized protein	0,85	2,77E-02
BL1611	abfA3	Alpha-L-arabinosidase	0,86	4,56E-02
BL1649		Hypothetical membrane protein possibly involved in transport	0,86	7,91E-06
BL0650		Uncharacterized protein	0,88	1,18E-04
BL1573		Probable glycogen operon protein GlgX	0,88	8,75E-05
BL1768		Probable LacI-type transcriptional regulator	0,89	3,53E-04
BL0905		Uncharacterized protein	0,89	2,97E-02
BL1658		Histidine kinase sensor of two component system	0,89	2,03E-05
BL0284		FemAB-like involved in interpeptide bridge formation in peptidoglycan	0,89	1,76E-03
BL0978	lacZ	Beta-galactosidase (EC 3.2.1.23) (Lactase)	0,91	1,01E-04
BL1304		Uncharacterized protein	0,91	3,41E-05
BL0028		Probable DEAD box-like helicase	0,91	5,92E-03
BL1363	gap	Glyceraldehyde 3-phosphate dehydrogenase C	0,91	2,77E-06
BL0056		Possible beta-hexosaminidase A	0,91	1,83E-03
BL0613		Probable integral membrane transporter	0,92	4,64E-07
BL0459		Flavodoxin-like domain-containing protein	0,92	9,28E-03
BL0152	aroP	Aromatic amino acid transport protein AroP	0,93	6,29E-05
BL0859	bex era	GTPase Era	0,95	4,93E-07
BL0631		Uncharacterized protein	0,95	3,12E-02
BL0936		AAA domain-containing protein	0,95	1,33E-02
BL0075		Permease of ABC transporter system for amino acids	0,96	2,96E-03
BL0860		Probable conserved integral membrane protein with CBS domain	0,96	4,82E-02
BL1246		DUF4190 domain-containing protein	0,97	3,81E-02

BL0305	rpsP	30S ribosomal protein S16	0,97	1,30E-02
BL0232	fabG	Probable 3-oxoacyl-[acyl-carrier protein] reductase	0,98	1,33E-02
BL1219	ptrB	Protease II	0,98	5,41E-03
BL1452		Conserved hypothetical transmembrane protein related to ComA	1,00	1,18E-02
BL0297	ilvN	Acetolactate synthase small subunit	1,00	2,27E-03
BL1159		Probable ABC transporter permease protein for peptides	1,00	3,14E-02
BL0183		Possible endo-1,5-alpha-L-arabinosidase	1,00	4,97E-02
BL1181		Peptidase C51 domain-containing protein	1,01	1,48E-04
BL1630	pgm	Phosphoglucomutase	1,01	5,48E-04
BL1654	lysS	Lysine--tRNA ligase (EC 6.1.1.6) (Lysyl-tRNA synthetase) (LysRS)	1,01	1,28E-02
BL1735		WYL domain-containing protein	1,02	9,69E-08
BL0621		Uncharacterized protein	1,02	2,11E-04
BL1555		Uncharacterized protein	1,03	8,10E-06
BL1329		Alpha-mannosidase	1,04	5,95E-04
BL0299		ATP binding protein of ABC transporter	1,05	6,25E-04
BL0670	nrdE	Ribonucleoside-diphosphate reductase (EC 1.17.4.1)	1,05	2,28E-03
BL0033	fruE	Fructose import binding protein FruE	1,05	2,81E-03
BL1670	trmB	tRNA (guanine-N(7)-)-methyltransferase (EC 2.1.1.33)	1,06	4,08E-02
BL1428	hrdB sigA	RNA polymerase sigma factor SigA	1,08	2,52E-05
BL1645		Response regulator of two-component system	1,08	6,03E-10
BL1526	agl	Oligo-1,6-glucosidase	1,08	2,10E-03
BL0048		Probable aminotransferase	1,10	2,95E-02
BL1665	thyA	Thymidylate synthase (TS) (TSase) (EC 2.1.1.45)	1,11	3,55E-02
BL0157		Uncharacterized protein	1,11	4,12E-05
BL0437		Probable TetR-type transcriptional regulator	1,11	3,92E-02
BL0194	kup2	Probable potassium transport system protein kup 2	1,11	4,82E-02
BL0440	zwf2 zwf	Glucose-6-phosphate 1-dehydrogenase (G6PD) (EC 1.1.1.49)	1,12	6,70E-03
BL0564		Cytosine-specific methyltransferase (EC 2.1.1.37)	1,15	3,52E-05
BL1164		Probable solute binding protein of ABC transporter system for sugars	1,17	3,90E-02
BL0076	gltL	ATP binding protein of ABC transporter for glutamate/aspartate	1,17	8,20E-03
BL0287		Nicotinate phosphoribosyltransferase (EC 6.3.4.21)	1,18	3,39E-02
BL1798	hup	DNA-binding protein HB1	1,18	9,25E-03
BL1531		Uncharacterized protein	1,18	2,59E-08
BL0272	araA	L-arabinose isomerase (EC 5.3.1.4)	1,22	6,07E-03
BL0310		Uncharacterized protein	1,23	6,37E-04
BL0332		MFS domain-containing protein	1,23	7,36E-10
BL0464		F5/8 type C domain-containing protein	1,24	1,97E-04
BL0230		Hypothetical transmembrane protein involved in EPS biosynthesis	1,24	1,41E-07
BL1623	rpiA	Ribose-5-phosphate isomerase A (EC 5.3.1.6)	1,25	1,44E-03
BL1516		Probable NhaP-type Na(+)/H(+) exchanger	1,25	1,02E-10
BL0856	pntB	NAD(P) transhydrogenase subunit beta (EC 7.1.1.1)	1,25	1,16E-05
BL0556	tgt	Queuine tRNA-ribosyltransferase (EC 2.4.2.29)	1,26	5,60E-03
BL1418	hpf	Ribosome hibernation promoting factor (HPF)	1,27	1,03E-02
BL0512		3-oxoacyl-[acyl-carrier protein] reductase	1,27	7,85E-09
BL1498		Peripla_BP_4 domain-containing protein	1,28	9,39E-03
BL0554		Probable cation-transporting ATPase V	1,30	4,01E-06
BL0123		Pseudouridine synthase (EC 5.4.99.-)	1,30	6,56E-04
BL0651		Conserved hypothetical membrane protein in MviN family	1,31	3,37E-06
BL1114		UmuC domain-containing protein	1,31	1,24E-04
BL1446		Thump_like domain-containing protein	1,32	5,76E-09
BL0198		FtsX domain-containing protein	1,35	3,01E-08
BL1534		Biotin transporter	1,36	2,72E-09
BL1226	trxA2	Thioredoxin	1,38	1,04E-03
BL0452		Possible permease protein of ABC transporter system	1,38	1,88E-02
BL1784		Uncharacterized protein	1,38	2,00E-02
BL0233		PS_pyruv_trans domain-containing protein	1,38	1,91E-03
BL1771	dcuC	C4-dicarboxylate transporter	1,39	5,14E-03
BL0129		Uncharacterized protein	1,39	3,70E-04
BL0187		Exo-alpha-L-arabinofuranosidase II	1,40	2,57E-02
BL0216		Glft2_N domain-containing protein	1,40	1,49E-10
BL1228		Uncharacterized protein	1,41	3,92E-09
BL0409	silP	Copper-transporting ATPase	1,46	1,06E-09
BL1320	mraY	Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13)	1,46	1,06E-03
BL1487		Uncharacterized protein	1,48	1,83E-03
BL1642		Putative desulfatase possibly for mucin	1,48	1,58E-07
BL0158		Uncharacterized protein	1,49	6,83E-12
BL0661		Uncharacterized protein	1,50	1,55E-10
BL0510	dedA	DedA integral membrane protein	1,51	3,21E-10
BL0783		Probable aminotransferase	1,51	2,07E-02
BL0543		LacI-type transcriptional regulator	1,51	6,11E-03
BL0156	aapA	Amino acid permease	1,52	5,42E-05
BL0125		Narrowly conserved hypothetical membrane protein	1,52	1,26E-06
BL0866	glgC	Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27)	1,53	1,49E-08
BL1782	hsdM	HsdM	1,53	9,03E-07
BL0672		Probable glycosyltransferase	1,55	1,13E-03

BL1265		DHO_dh domain-containing protein	1,55	3,95E-08
BL0153		Probable TetR-like transcriptional regulator	1,56	2,50E-02
BL0040		TM2 domain-containing protein	1,56	1,98E-03
BL1288	nusG	Transcription termination/antitermination protein NusG	1,56	3,51E-03
BL1752	nrdD	Anaerobic ribonucleoside-triphosphate reductase	1,57	8,24E-14
BL1266	nox	NADH oxidase	1,57	2,21E-05
BL1532		Uncharacterized protein	1,58	2,93E-20
BL0285		dITP/XTP pyrophosphatase (EC 3.6.1.66)	1,58	8,10E-06
BL0031		Histidine kinase sensor of two-component system	1,58	2,27E-02
BL1210	galK	Galactokinase	1,59	3,58E-06
BL0189		Sugar permease of ABC transporter system	1,61	1,43E-03
BL0021	gluA	ATP-binding protein of ABC transporter for glutamate	1,61	7,15E-04
BL1753	nrdG	Anaerobic ribonucleoside-triphosphate reductase activating protein	1,63	1,21E-03
BL0208		Transport permease protein	1,64	2,64E-13
BL0138		Uncharacterized protein	1,69	2,83E-19
BL1566		Conserved hypothetical transmembrane protein with duf6	1,69	2,90E-04
BL1358		Uncharacterized protein	1,71	2,19E-05
BL0207	rgpDc	Protein possibly involved in ATP-driven polysaccharide export	1,72	7,11E-15
BL0012		Uncharacterized protein	1,72	1,60E-05
BL1641		Uncharacterized protein	1,72	1,50E-10
BL1624		Ribonuclease H (EC 3.1.26.4)	1,72	4,42E-04
BL1761		Probable beta-1,3-exoglucanase	1,74	2,81E-04
BL0228		dTDP-4-dehydrorhamnose reductase (EC 1.1.1.133)	1,76	1,01E-03
BL1212		Probable DeoR-type transcriptional regulator	1,79	3,50E-12
BL0165		Possible symporter	1,79	3,32E-07
BL0037		Probable efflux-type transporter	1,80	2,09E-05
BL1314		DNA helicase (EC 3.6.4.12)	1,81	1,54E-14
BL0426		LacI-type transcriptional regulator	1,81	3,84E-14
BL1081		Possible endonuclease III	1,83	4,97E-03
BL0276		Ribonuclease (EC 3.1.26.4)	1,84	2,48E-06
BL1357	sigH	RNA polymerase sigma-E factor	1,84	7,73E-04
BL0660		Widely conserved protein with eukaryotic protein kinase domain	1,90	4,95E-10
BL0615	ahpC	Alkyl hydroperoxide reductase C (EC 1.11.1.15)	1,91	1,34E-16
BL1364		TPK_B1_binding domain-containing protein	1,91	1,52E-07
BL1770		IU_nuc_hydro domain-containing protein	1,91	8,51E-03
BL0628		Aminotransferase (EC 2.6.1.-)	1,91	4,53E-05
BL0146		Possible arabinosidase	1,92	2,98E-06
BL0274		Probable sugar kinase	1,93	3,15E-03
BL1538		Protein with similarity to holo-[acyl-carrier protein] synthase	1,94	5,55E-07
BL0671	nrdF	Ribonucleoside-diphosphate reductase subunit beta (EC 1.17.4.1)	1,97	9,42E-03
BL0467		Uncharacterized protein	1,98	3,13E-21
BL1142		L-asparaginase	1,99	5,72E-05
BL0646	rsmG	Ribosomal RNA small subunit methyltransferase G (EC 2.1.1.-)	1,99	5,80E-12
BL1702		Hypothetical membrane protein possibly involved in transport	2,00	7,32E-18
BL1359		Uncharacterized protein	2,01	2,16E-03
BL0008		Uncharacterized protein	2,02	3,10E-25
BL1808		Uncharacterized protein	2,02	1,18E-04
BL1666	dfrA	Dihydrofolate reductase	2,03	5,70E-09
BL1456	rimI	Probable ribosomal-protein-alanine N-acetyltransferase	2,05	3,08E-16
BL0385		Holin-like protein similar to orf of bacteriophage BK5-T	2,05	5,89E-04
BL1652		AraJ-like protein probably involved in transport of arabinose polymers	2,06	3,76E-03
BL0279	pgi gpi	Glucose-6-phosphate isomerase (GPI) (EC 5.3.1.9)	2,08	1,34E-22
BL1457	tsaD gcp	tRNA N6-adenosine threonylcarbamoyltransferase (EC 2.3.1.234)	2,09	2,68E-06
BL1305	hrpA	ATP-dependent helicase	2,10	8,45E-27
BL0137	pip	ABC2_membrane domain-containing protein	2,12	3,90E-35
BL0446		Hypothetical membrane protein with similarity to phage infection protein	2,12	6,41E-07
BL1211	galT1	Galactose-1-phosphate uridylyltransferase	2,16	5,42E-22
BL0562		UPF0145 protein BL0562	2,17	4,19E-04
BL0507		Uncharacterized protein	2,20	1,71E-03
BL0227	rmlA	Glucose-1-phosphate thymidylyltransferase (EC 2.7.7.24)	2,23	5,73E-05
BL0625		Uncharacterized protein	2,24	1,42E-14
BL0550	fba	Fructose-bisphosphate aldolase (FBP aldolase) (EC 4.1.2.13)	2,24	1,74E-12
BL0558		Uncharacterized protein	2,24	1,70E-07
BL1090		Possible alcohol dehydrogenase	2,25	2,99E-09
BL0627		Uncharacterized protein	2,33	1,22E-07
BL0977		Uncharacterized protein	2,34	3,34E-12
BL0626		Uncharacterized protein	2,35	5,90E-20
BL0561		N-acetyltransferase domain-containing protein	2,36	4,02E-04
BL0139		Possible NADPH-flavin oxidoreductase flavin reductase	2,36	9,62E-18
BL0457		NAD(P)-bd_dom domain-containing protein	2,49	3,11E-20
BL0624		Uncharacterized protein	2,52	2,05E-23
BL1710		Possible XylR-type repressor	2,56	1,19E-05
BL0805		Uncharacterized protein	2,57	3,10E-03
BL0200		Acyl_transf_3 domain-containing protein	2,61	1,22E-30
BL1574		Probable repressor in the Rok (NagC/XylR) family	2,63	2,40E-07

BL0122	lspA	Lipoprotein signal peptidase (EC 3.4.23.36)	2,63	2,27E-03
BL0368		Uncharacterized protein	2,65	2,43E-05
BL1360		Uncharacterized protein	2,68	1,17E-04
BL0155		Large transmembrane protein possibly involved in transport	2,68	3,90E-49
BL0566		Glyco_trans_2-like domain-containing protein	2,70	7,37E-14
BL0370		Uncharacterized protein	2,78	3,84E-06
BL1643	galT2 galT	Galactose-1-phosphate uridylyltransferase (EC 2.7.7.12)	2,79	2,72E-07
BL1667	ptpA	Low molecular weight protein-tyrosine-phosphatase	3,04	1,51E-16
BL0950	pflA	Pyruvate formate-lyase-activating enzyme (EC 1.97.1.4)	3,08	3,29E-16
BL0560		Uncharacterized protein	3,11	4,66E-21
BL1674		Probable glycosyltransferase	3,23	2,38E-55
BL1521	msmE	Sugar binding protein of ABC transporter system	3,24	1,91E-05
BL0951	pfl	Formate acetyltransferase	3,25	1,24E-64
BL1132	dxs	1-deoxy-D-xylulose-5-phosphate synthase (EC 2.2.1.7)	3,29	2,75E-58
BL1671	galE2	UDP-glucose 4-epimerase	3,32	1,96E-11
BL1414		Possible DNA binding protein	3,35	7,35E-07
BL1524		Uncharacterized protein	3,45	9,68E-23
BL1134		Uncharacterized protein	3,63	1,21E-06
BL1135		ATP binding protein of ABC transporter	3,66	1,12E-32
BL0567		Uncharacterized protein	3,91	2,72E-26
BL0835		Uncharacterized protein	4,80	1,03E-56
BL1695		ATP binding protein of ABC transporter for pentoses	5,96	1,06E-67
BL1694		Probable sugar binding protein of ABC transporter for pentoses	6,34	2,57E-79
BL1697		Uncharacterized protein	7,47	4,21E-153

Table S 2: List of *B. longum* NCC 2705 genes enriched in the glucose grown cells and significatively associated to different Gene Onthologies (GO) upon GO enrichement analysis using the BINGO plugin of Cytoscape.

GO-ID	p-value	n° genes	GO description	Genes in test set
15672	3,08E-03	5	monovalent inorganic cation transport	BL0359|BL0358|BL0362|BL0361|BL1046
6754	5,60E-03	4	ATP biosynthetic process	BL0359|BL0358|BL0362|BL0361
15985	5,60E-03	4	energy coupled proton transport, down electrochemical gradient	BL0359|BL0358|BL0362|BL0361
15986	5,60E-03	4	ATP synthesis coupled proton transport	BL0359|BL0358|BL0362|BL0361
15992	5,60E-03	4	proton transport	BL0359|BL0358|BL0362|BL0361
6818	5,60E-03	4	hydrogen transport	BL0359|BL0358|BL0362|BL0361
9145	5,60E-03	4	purine nucleoside triphosphate biosynthetic process	BL0359|BL0358|BL0362|BL0361
46034	5,60E-03	4	ATP metabolic process	BL0359|BL0358|BL0362|BL0361
6119	5,60E-03	4	oxidative phosphorylation	BL0359|BL0358|BL0362|BL0361
9205	5,60E-03	4	purine ribonucleoside triphosphate metabolic process	BL0359|BL0358|BL0362|BL0361
9206	5,60E-03	4	purine ribonucleoside triphosphate biosynthetic process	BL0359|BL0358|BL0362|BL0361
6812	7,24E-03	5	cation transport	BL0359|BL0358|BL0362|BL0361|BL1046
9144	9,29E-03	4	purine nucleoside triphosphate metabolic process	BL0359|BL0358|BL0362|BL0361
6811	9,64E-03	6	ion transport	BL1657|BL0359|BL0358|BL0362|BL0361 |BL1046
6432	1,08E-02	2	phenylalanyl-tRNA aminoacylation	BL1066|BL1067
16310	1,43E-02	4	phosphorylation	BL0359|BL0358|BL0362|BL0361
9199	1,43E-02	4	ribonucleoside triphosphate metabolic process	BL0359|BL0358|BL0362|BL0361
9201	1,43E-02	4	ribonucleoside triphosphate biosynthetic process	BL0359|BL0358|BL0362|BL0361
6793	1,43E-02	5	phosphorus metabolic process	BL0359|BL0721|BL0358|BL0362|BL0361
6796	1,43E-02	5	phosphate metabolic process	BL0359|BL0721|BL0358|BL0362|BL0361
34220	2,06E-02	4	ion transmembrane transport	BL0359|BL0358|BL0362|BL0361
6633	3,02E-02	2	fatty acid biosynthetic process	BL1536|BL1537
50794	3,28E-02	15	regulation of cellular process	BL0151|BL1291|BL1356|BL1312|BL1267|BL1633| BL1336|BL1530|BL1398|BL0275|BL0846|BL1657| BL1613|BL0721|BL0930
9142	3,80E-02	4	nucleoside triphosphate biosynthetic process	BL0359|BL0358|BL0362|BL0361
50789	3,95E-02	15	regulation of biological process	BL0151|BL1291|BL1356|BL1312|BL1267|BL1633| BL1336|BL1530|BL1398|BL0275|BL0846|BL1657| BL1613|BL0721|BL0930
31323	4,04E-02	12	regulation of cellular metabolic process	BL0846|BL0151|BL1657|BL1613|BL0930|BL1291| BL1312|BL1633|BL1336|BL1530|BL1398|BL0275
19222	4,04E-02	12	regulation of metabolic process	BL0846|BL0151|BL1657|BL1613|BL0930|BL1291| BL1312|BL1633|BL1336|BL1530|BL1398|BL0275
9059	4,62E-02	21	macromolecule biosynthetic process	BL0071|BL0170|BL1291|BL1356|BL0476|BL1565| BL1267|BL1355|BL1578|BL1099|BL1066|BL1594| BL1186|BL1571|BL1354|BL1067|BL1100|BL0403| BL0634|BL0721|BL0719

Table S 3: List of *B. longum* NCC 2705 genes enriched in the galactose grown cells and significatively associated to different Gene Onthologies (GO) upon GO enrichement analysis using the BINGO plugin of Cytoscape.

GO-ID	p-value	n° genes	GO description	Genes in test set
5975	1,14E-07	35	carbohydrate metabolic process	BL0272|BL1360|BL1543|BL1322|BL0951|BL1643|BL0279|BL0274|BL1363|BL1761|BL0056|BL0550|BL1320|BL0715|BL0978|BL0739|BL1623|BL0999|BL1329|BL1526|BL1691|BL0183|BL1630|BL1611|BL1671|BL1211|BL0440|BL0561|BL0187|BL1210|BL1573|BL0227|BL1359|BL0866|BL0228
19318	2,79E-07	17	hexose metabolic process	BL1691|BL1643|BL0279|BL1363|BL1671|BL1211|BL0440|BL0550|BL1210|BL1573|BL0715|BL0739|BL1623|BL0999|BL1329|BL0866|BL0228
5996	1,29E-05	20	monosaccharide metabolic process	BL0272|BL1691|BL1611|BL1643|BL0279|BL1363|BL1671|BL1211|BL0440|BL0561|BL0550|BL1210|BL1573|BL0715|BL0739|BL1623|BL0999|BL1329|BL0866|BL0228
6066	2,36E-05	20	alcohol metabolic process	BL0272|BL1691|BL1611|BL1643|BL0279|BL1363|BL1671|BL1211|BL0440|BL0561|BL0550|BL1210|BL1573|BL0715|BL0739|BL1623|BL0999|BL1329|BL0866|BL0228
44262	2,53E-05	21	cellular carbohydrate metabolic process	BL0272|BL1691|BL1611|BL1643|BL0279|BL1363|BL1671|BL1211|BL0440|BL0561|BL0550|BL1210|BL1573|BL0715|BL0739|BL0227|BL1623|BL0999|BL1329|BL0866|BL0228
6006	5,54E-05	11	glucose metabolic process	BL0715|BL0739|BL1691|BL1623|BL0999|BL0866|BL0279|BL1363|BL0440|BL0550|BL1573
16052	1,54E-03	10	carbohydrate catabolic process	BL0715|BL0272|BL1691|BL1623|BL1543|BL0279|BL1211|BL0440|BL0550|BL1573
44275	2,44E-03	9	cellular carbohydrate catabolic process	BL0715|BL0272|BL1691|BL1623|BL0279|BL1211|BL0440|BL0550|BL1573
46365	3,85E-03	8	monosaccharide catabolic process	BL0715|BL0272|BL1691|BL1623|BL0279|BL1211|BL0440|BL0550
19320	4,17E-03	7	hexose catabolic process	BL0715|BL1691|BL1623|BL0279|BL1211|BL0440|BL0550
46164	5,38E-03	8	alcohol catabolic process	BL0715|BL0272|BL1691|BL1623|BL0279|BL1211|BL0440|BL0550
6012	8,03E-03	4	galactose metabolic process	BL1643|BL1671|BL1211|BL1210
5977	8,72E-03	3	glycogen metabolic process	BL0999|BL0866|BL1573
6112	8,72E-03	3	energy reserve metabolic process	BL0999|BL0866|BL1573
6007	1,36E-02	6	glucose catabolic process	BL0715|BL1691|BL1623|BL0279|BL0440|BL0550
45454	1,44E-02	4	cell redox homeostasis	BL0615|BL0821|BL1226|BL1266
5978	1,82E-02	2	glycogen biosynthetic process	BL0999|BL0866
44282	2,03E-02	9	small molecule catabolic process	BL0715|BL0272|BL1691|BL1623|BL0279|BL0285|BL1211|BL0440|BL0550
34637	2,51E-02	6	cellular carbohydrate biosynthetic process	BL0227|BL0999|BL0866|BL0228|BL0279|BL0561
44264	3,48E-02	4	cellular polysaccharide metabolic process	BL0227|BL0999|BL0866|BL1573
19725	3,48E-02	4	cellular homeostasis	BL0615|BL0821|BL1226|BL1266
44042	3,55E-02	3	glucan metabolic process	BL0999|BL0866|BL1573
6740	3,55E-02	3	NADPH regeneration	BL0715|BL1623|BL0440
6073	3,55E-02	3	cellular glucan metabolic process	BL0999|BL0866|BL1573
6098	3,55E-02	3	pentose-phosphate shunt	BL0715|BL1623|BL0440

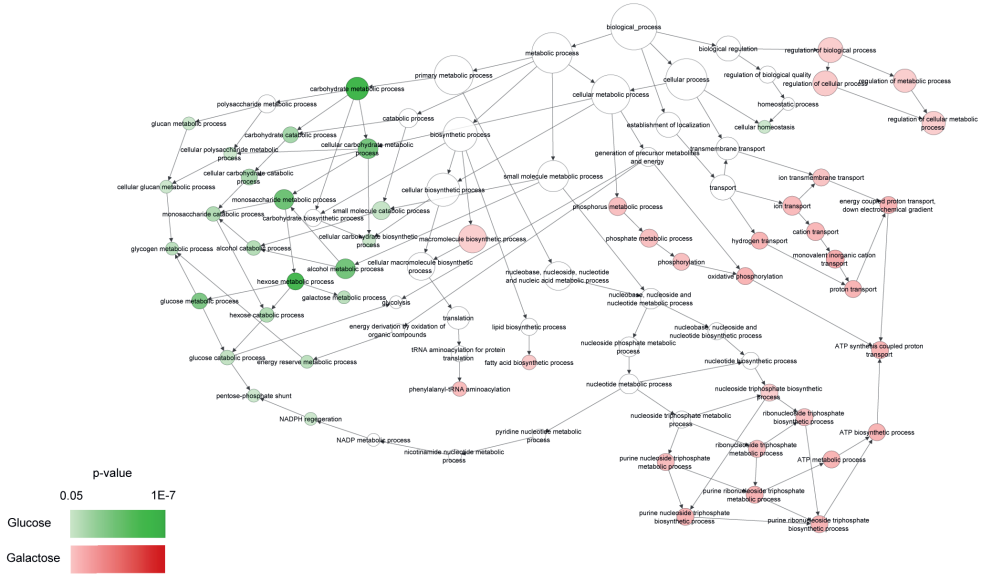


Figure S 1: Gene Ontology (GO) functions enriched in glucose or galactose upregulated genes (FDR p-value <0.05). GO enrichment analysis was performed using the Cytoscape BINGO plugin, using the hypergeometric statistical test. Enriched functions (p-value <0.05) in the glucose condition are displayed in red scale while those enriched (p-value<0.05) in the galactose condition are presented in green scale.

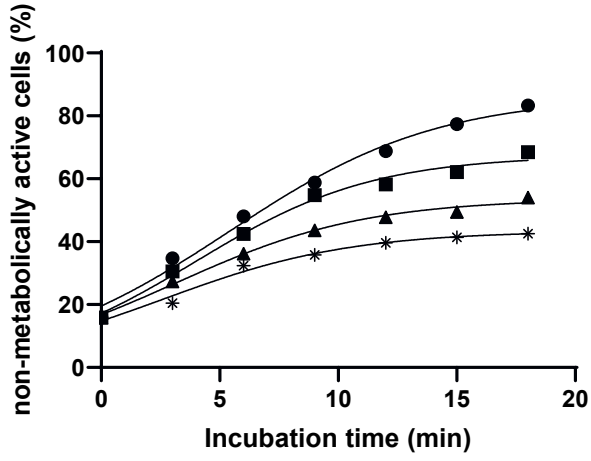


Figure S 2: Metabolic activity of *B. longum* NCC 2705 stationary grown glucose cells exposed to pH 2.9 (circles), pH 3.0 (squares), pH 3.2 (triangles) and pH 3.5 (stars) for a duration of 18 minutes at 37°C. Esterase activity (cFDA staining) was quantified by flow cytometry and used as a measure of metabolic activity in the tested cultures.

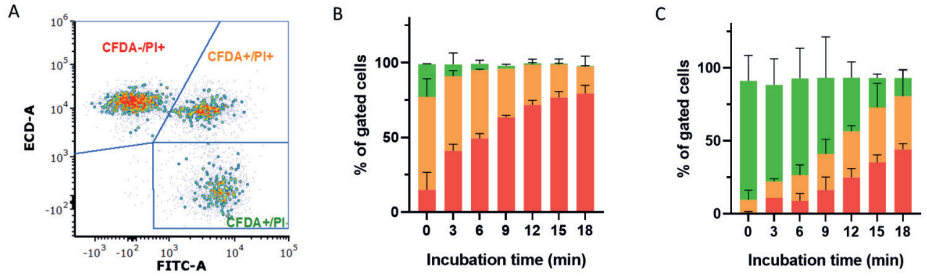


Figure S 3: *B. longum* NCC 2705 population distribution upon exposure to a pH of 2.9. Panel A represents the population distribution of a galactose stationary grown culture exposed to a pH of 2.9 for 12 minutes. Population distribution was evaluated using a double labelling strategy: propidium iodide (PI) to evaluate membrane integrity (distributed on the Y, ECD-A axis) and CFDA to evaluate metabolic activity (distributed on the X, FITC-A axis). Distribution of metabolically active & non permeable cells (CFDA+/PI-, green), metabolically active & permeable cells (CFDA+/PI+, orange) and non-metabolically permeable cells (CFDA-/PI+) within stationary glucose (Panel B) and galactose (Panel B) grown cultures exposed to a pH of 2.9 is further depicted.

CHAPTER 7

7

A sRNA encoded in the 16S-23S intergenic region of *B. longum* NCC 2705 plays a global regulatory role, including regulation of serpin production

S. Duboux [1,2] | B. Bogicevic [1] | J.A. Muller [1] | J. P. Tan [1] | S. Pruvost [1] | C. Joyce [1] | A. Mercenier [2] | M. Kleerebezem [2]

[1] Nestlé Research, Route du Jorat 57, CH 1000 Lausanne 26, Switzerland
[2] Host-Microbe Interactomics Group, Wageningen University & Research, De Elst 1, 6708WD Wageningen, The Netherlands

Abstract

The serine protease inhibitor (serpin; *BL0108*) of *B. longum* NCC 2705 was previously demonstrated to be essential for the strain's immunoprotective effect in a murine model of celiac disease [64]. We have also reported that serpin production was inhibited by the presence of glucose and induced by galactose and fructose [140], although the underlying regulation mechanism remained to be deciphered. In the present study, we first highlight strong discrepancies between the serpin transcript and protein levels in NCC 2705 grown on glucose or galactose, suggesting post-transcriptional regulation possibly mediated by a regulatory sRNA. Using transcriptomics, we identified and quantified three sRNA that could inhibit the translation of the serpin mRNA by base-pairing (with energy of interaction < -16 kcal/mol). Out of those, sBLO-161 was the only one to be consistently differentially expressed in glucose and galactose grown cultures (2.8 fold increase in glucose). The sBLO-161 sRNA, located between the 16S and 23S rDNA sequences, was not only predicted to stably bind to the *BL0108* mRNA by its 5' and 3' ends, but also was foreseen to form two internal stem-loops that could inhibit translation of the *BL0108* mRNA. Strikingly, this sRNA was predicted to form highly similar RNA-duplex structures with a range of other mRNAs in NCC 2705 that encode various important functions, like stress response, cell wall synthesis, carbohydrate transport and metabolism. We were particularly intrigued by the predicted interaction of sBLO-161 with the mRNA encoding the xylose/fructose-6-phophoketolase in this strain, an enzyme that plays a pivotal role in the central energy metabolism of bifidobacteria. Finally, we show that sBLO-161 is relatively well conserved in the *B. longum* subspecies, although the functions it regulates may vary between subspecies. For example, in *B. longum* subsp. *infantis* ATCC 15697, the sBLO-161 sRNA homologue is predicted to not participate in serpin regulation while retaining its role in regulating translation of genes implicated in the stress response or carbohydrate metabolism. Overall, we describe the first bifidobacterial sRNA that is predicted to exert a global regulatory function, not only controlling production of serpin in *B. longum* NCC 2705, but also exerting post-transcriptional control of a wide range of important functions in this strain.

Introduction

Strains belonging to the *Bifidobacterium* genus are among the first bacteria to establish in the human gastrointestinal tract (GIT) where they are particularly dominant during the breastfeeding period [7]. Their presence in early infancy is proposed to be critical for the newborn's immune system maturation and several studies have linked their absence or low abundance to later in life immune disorders such as atopy [38] or asthma [39].

B. longum subsp. *longum* NCC 2705 (hereafter *B. longum* NCC 2705) was originally isolated from the feces of a healthy infant and was the first *Bifidobacterium* strain for which the complete genome sequence was published [16]. The genome is particularly rich in genes linked to carbohydrates metabolism, which provides most of the strain required energy. This is reflected by the large repertoire of carbohydrate degrading enzymes encoded in the genome [16,19], as well as by several additional features that support its gastrointestinal lifestyle [26,54,57]. Amongst these, *B. longum* NCC 2705 produces a serine protease inhibitor (serpin) encoded by the *BL0108* gene, which forms covalent products with pancreatic and neutrophil elastases thereby inhibiting their function [61]. The serpin protein was proposed to protect bidifobacteria against host-derived proteases and hence to provide them a survival advantage in the competitive intestinal environment [61,137]. The serpin's capacity to inhibit the human neutrophil elastase may also be involved in the immunomodulatory capacities of the strain [63] as elastase is released by activated neutrophils at the sites of intestinal inflammation [65]. Accordingly, serpin was demonstrated to play a pivotal role in the anti-inflammatory effect of *B. longum* NCC 2705 in a mouse model of gluten sensitivity [64]. We have shown previously that serpin regulation in this strain was different from that of *B. breve* where serpin was shown to be regulated by a protease activated two-component system [139]. In *B. longum* NCC 2705, serpin production is controlled by carbohydrates, i.e. galactose and fructose induce its production while glucose represses it [140]. We suggested that the serpin promoter might be controlled by AraQ [140], as a sequence presenting a modest resemblance to the *cis*-element of this (proposed) global regulator [181] was detected upstream of the *BL0108* gene. However, the mechanism regulating serpin production in *B. longum* NCC 2705 warrants further investigation as this protein corresponds to a key effector molecule mediating beneficial health effects.

In bacteria, small non-coding RNAs (sRNAs) are proposed to allow cells to switch rapidly between different physiological states [259] and they were, for example, reported to be implicated in the regulation of bacterial stress response [260]. Bacterial sRNAs typically range from 50 to 350 nucleotides in length, are highly structured and one of the mechanisms proposed for their regulatory role is by modulation of mRNAs translation through a base-pairing mechanism [261-263]. Two principal categories of bacterial sRNAs affecting mRNA translation have been defined: i) cis-antisense sRNAs, expressed from the strand opposite to their target gene [264] and (ii) trans-acting sRNAs, expressed either from intergenic regions [265] or from untranslated regions close to a coding sequence (CDS) [266]. Different tools enable the detection of bacterial sRNAs in transcriptomics datasets. The APERO software showed an improved performance and accuracy compared to several other available tools, which was explained by its sRNA detection approach that focusses on the localization of the ends of the transcripts [267]. Furthermore, different models are available to predict RNA-RNA interactions based on the stability of the base-pairing (minimum free energy method) and the accessibility of the binding-region upon folding [268]. Using these basic predictions, tools have been developed to specifically predict bacterial whole-genome sRNA targets and analyse their detailed interactions [269,270].

Research aiming at deciphering the regulatory role of bifidobacterial sRNA is relatively scarce. A study conducted in *B. animalis* subsp. *lactis* KDLS 2.0603 identified eleven differentially expressed sRNAs possibly involved in the regulation of the strain's gastrointestinal-like stress response. The authors also performed a sRNA prediction in NCC 2705 and detected the presence of 39 sRNA candidates in this strain, although without investigating their potential role [271]. Bottacini et al. identified three sRNAs in *B. breve* UCC2003 and suggested that they are implicated in the regulation of riboflavin and thiamin production as well as in the transport of metal ions [272].

With the present study, we demonstrate that the production of serpin is not controlled by differential gene transcription in glucose versus galactose grown cells, suggesting a post-transcriptional regulation mechanism. To investigate the possible involvement of sRNA in the regulation of serpin production we mined paired-end RNA sequencing data obtained from glucose or galactose grown cells for candidate sRNA molecules. Leveraging the available RNA-RNA interaction predictions tools, we identified three sRNA that have the capacity to bind to the *BL0108* encoded mRNA and evaluated

their expression profile in both glucose and galactose grown cultures. The sRNA having the strongest *BL0108* binding capacity is located in the 16S-23S rDNA intergenic region and was predicted to not only target the serpin transcript, but to also play a global regulatory role in NCC 2705 and other *B. longum* subspecies. Our results highlight the importance of this sRNA in the regulation of key physiological features of *B. longum* NCC 2705, influenced by the carbon source used for growth.

Results

Growth on galactose induces discrepant serpin protein and mRNA levels. To gain insight in the potential regulation mechanisms of *Bl0108*, the serpin protein & RNA levels obtained upon growth on glucose or galactose as sole carbon source were compared. Using a serpin specific ELISA, we confirmed that glucose or galactose mid-exponentially grown cultures produced distinct serpin protein levels (291 and 5232 pg / mg total protein, respectively) (Figure 1 A). However, RNAseq data obtained from the same cultures [273] revealed no significant difference in the *BL0108* mRNA levels when comparing the glucose and galactose grown cells (i.e., 7.8 vs 7.9 log2 counts, respectively) (Figure 1 B). Furthermore, and in contrary to the AraQ controlled promoter P_{BL1359}, the previously predicted serpin promoter (P_{BL0109}) [137] fused to a mCherry protein was not activated upon growth of the strain on galactose (Figure 1 C). As opposed to our initial hypothesis [140], these results advocate that P_{BL0109} does not seem to be controlled by AraQ and highlight a discrepancy between the serpin protein and mRNA levels, suggesting that serpin production may be controlled through a post-transcriptional regulation mechanism.

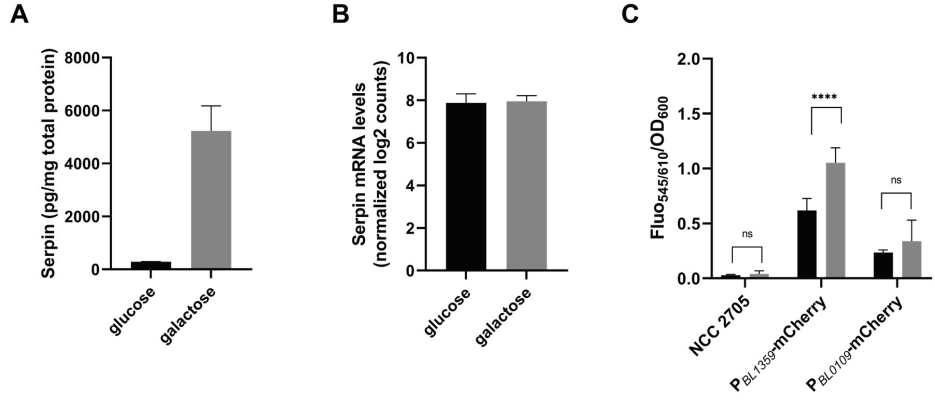


Figure 1: Serpin (BL0108) protein (panel A) and RNA (panel B) levels and promoter activation (panel C) in mid-exponentially glucose or galactose grown *B. longum* NCC 2705. Serpin protein levels were measured using a serpin-specific ELISA (panel A), while RNA levels were retrieved from previously obtained RNAseq data (panel B). Panel C shows the activation levels of the serpin promoter (P_{BL0109}) and an AraQ controlled promoter (P_{BL1359}) upon growth on glucose (black bars) or galactose (grey bars). Fluorescence levels of *B. longum* NCC 2705 derivatives harboring a fusion of each of the two promoters with the mCherry fluorescent proteins (respectively P_{BL1359} & P_{BL0109}) were evaluated at 545 & 610 nm (excitation & emission wavelenghts) upon growth on glucose or galactose as sole carbon source. Statistical analysis was performed using ANOVA followed by a Sidak's multiple test (ns>0.05; ****<0.001).

Several sRNAs are predicted to interact with the *BL0108* mRNA sequence. To investigate the potential involvement of sRNA in the post-transcriptional control of *BL0108*, we performed an initial sRNA detection in a mid-exponential glucose grown culture of *B. longum* NCC 2705. All detected transcripts were filtered on size (between 50-350 bp) and frequency of detection (>50) to obtain a list of 198 candidate sRNA transcripts (Table S 1) that were classified according to their localization relative to a coding region as proposed by Leonard and colleagues [267] (Figure S 1 A). Fifty (50) candidate sRNA corresponded to potentially degraded coding mRNA regions (classified "I") and were discarded. Of the remaining 148 candidate sRNAs, 106 transcripts were derived from intergenic regions (classified as "P", "O" or "Div"), and 42 transcripts overlapped the 5' or 3' ends of genes (classified as "3'UTR", "3'Ai" or "5' UTR") (Figure S 1 B).

The potential interaction of the 148 candidate sRNA with the *BL0108* transcript were analysed using RNApredator, revealing that 5 of the predicted sRNA could stably bind the *BL0108* mRNA sequence with a free-energy of interaction <-16 kcal / mol (Table S 2, includes sequences), which is a threshold that has previously been employed to predict sRNA targets [274]. Three of these candidate sRNA were selected because they were detected in all biological replicates with a relatively consistent size (<350 bp and less than 1.5-fold variation in size).

The first candidate sRNA was encoded in the intergenic region between the 16S and 23S rDNA coding genes of the *B. longum* NCC 2705 genome. This strain's genome harbors four 16S and 23S rDNA encoding loci, which all encompass the detected candidate sRNA sequence, hampering the exact definition of its origin. Therefore, a name for this sRNA was arbitrarily assigned according to the first copy of the 16S-23S rDNA region found within the NCC 2705 genome, whereby the 5'-end of the sRNA is located at position 160871 and therefore named with the identifier sBLO-161 according to the nomenclature system proposed by Li and colleagues [275]. APERO detected the sBLO-161 presence in all biological replicates with a consistent size (168-169 bp). This sRNA was predicted to have a particularly strong affinity to the *BL0108* mRNA, as the free energy associated with its hybridization to this target was the lowest (-44.5 kcal/mol), whereas the free energy associated with the unfolding of both the mRNA and sRNA was amongst the highest (10.4 and 12.84 kcal/mol, respectively) (Table 1). The second candidate sRNA was encoded within the *B. longum* NCC 2705

genome overlapping the *groES* encoding gene (5' position at 1968821; designated sBLO-1969) was also detected in all samples with a consistent size (147-148 bp). Its free energy of binding and unfolding to the *BL0108* mRNA was lower than the ones found for sBLO-161 (Table 1). The third candidate sRNA, for which the sequence overlaps the *trpS* encoding gene (5' position at 15777901; designated sBLO-1578) displayed considerable variation in size (from 263 to 336 bp) in the samples, which may be due to the sequencing library preparation procedure that fragmented cDNA to a target size of around 200 bp, and therefore suggests that this variably sized candidate sRNA is probably derived from a larger RNA molecule (Table 1). We thus identified sBLO-161 as the best serpin regulatory sRNA candidate.

Table 1: List of candidate sRNA interacting with the *BL0108* mRNA as predicted by RNApredator. The 5'-end position, class, length and genomic context detected by APERO for each of these sRNA is described. The energy of interaction with the serpin transcript as predicted by RNApredator (kcal/mol), calculated as the sum of the hybridization energy gained by the formation of the duplex RNA and the unfolding energies for both mRNA and sRNA are detailed within parenthesis. Finally, the expression profile (fold change and adjusted p-value) detected by APERO is indicated for each candidate sRNA.

	Genome position of the sRNA 5' end (strand)	**Class**	**sRNA length (bp)**	**Genomic context**	**energy of interaction (kJ mol^{-1})**	**glu:gal expression (fold; adj. p-value)**
sBLO-161	160871 (+)	O	168-9	Intergenic between 16S and 23S rRNA genes	-21.27 (-44.5; 10.4; 12.84)	2.77; 0.003
sBLO-1969	1968821 (+)	5U	147-148	Overlaps with start of *groES* (*BL1558*)	-16.63 (-32.1; 8.17; 7.3)	1.90; 0.056
sBLO-1578	15777901-74 (+)	5U	263-336	Overlaps with start of *trpS* (*BL0600*)	-16.95 (-45.7; 15.01; 13.74)	2.49; 0.125

sBLO-161 is upregulated in glucose. To further support a role of the identified sRNAs in the post-transcriptional regulation of *BL0108*, we investigated their expression profile in glucose or galactose grown *B. longum* NCC 2705. The complete RNAseq dataset (biological triplicate measurements in cells logarithmically growing in galactose or glucose medium) were subjected to the APERO sRNA detection pipeline, and the read counts specific to the three sRNA (sBLO-161, sBLO-1969, sBLO-1578) were retrieved. For sBLO-161, the read counts mapped to the different 16S-23 rDNA intergenic regions were summed to assess sBLO-161 expression. Notably, sBLO-1578 was not differentially expressed between the two conditions. The sBLO-1969 sRNA displayed the highest expression level but only showed a trend to be lower expressed in galactose growing cells as its expression values appeared to be somewhat variable (10408±993 in glucose, 5469±1913 in galactose). The sBLO-161 sRNA was less abundant than sBLO-1969 but displayed a very consistent differential expression level when comparing the two growth conditions (3975±343 in glucose; 1435±200 in galactose; Figure 2), suggesting a possible role for this sRNA in the inhibition of *BL0108* translation when the NCC 2705 is grown on glucose as sole carbon source. The combination of the weaker binding profiles of sBLO-1969 and sBLO-1578 and their lack of clear glucose or galactose mediated expression, led us to discard these candidate sRNA for a putative role in serpin regulation.

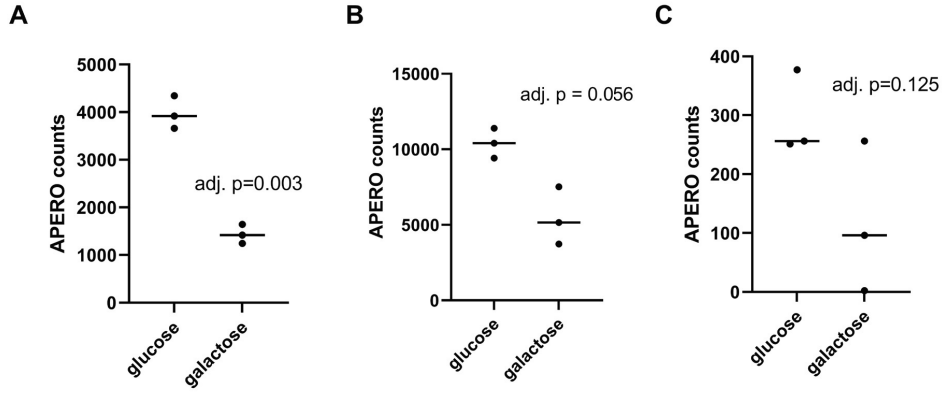


Figure 2: Expression profiles of detected sRNA targeting *BL0108* in *B. longum* NCC 2705. Detection of sRNA was performed with APERO on RNAseq data derived from the different replicates of the strain grown on galactose or glucose as sole carbon source. For sBLO-161, counts represent the sum of all reads found through the different 16S-23 rDNA intergenic regions (panel A). Specific counts were as well retrieved for sBLO-1969 (panel B) and sBLO-1578 (panel C) in each condition. Statistical analysis was performed using a multiple unpaired t-test with Welsh multiplicity correction.

sBLO-161 is predicted to efficiently interact with *BL0108* mRNA. To confirm the above suggested sBLO-161 role in the post-transcriptional regulation of the *BL0108* transcript, we further investigated the predicted interaction of this sRNA with the target *BL0108* mRNA. We could confirm using the intaRNA platform [270] that sBLO-161 binds to several positions within the *BL0108* mRNA sequence (Figure S 2 A). Analysis revealed a remarkable binding profile in the first 240 nucleotides of the *BL0108* mRNA, as both the 5'- (nt 09-36) and 3'- (nt 136-161) ends of sBLO-161 (Figure S 2 B) could bind the target mRNA region within a range of 80 bp (Table S 3). Furthermore, the sBLO-161 sRNA sequence located between those two binding sites (nt 58-144) was predicted to form a two-stem-loop structure with a strong affinity (free energy of -48.41 kcal/mol; Figure S 2 C). We propose that the RNA duplex structure formed (Figure 3) is suppressing translation of the serpin transcript. As sBLO-161 is 2.8 fold more abundant in glucose (as compared to galactose) grown NCC 2705, this may explain the lower levels of serpin harbored by the glucose grown cells. Overall, these results provide a strong support for a role of sBLO-161 in post-transcriptional regulation of serpin production (*BL0108*) in *B. longum* NCC 2705.

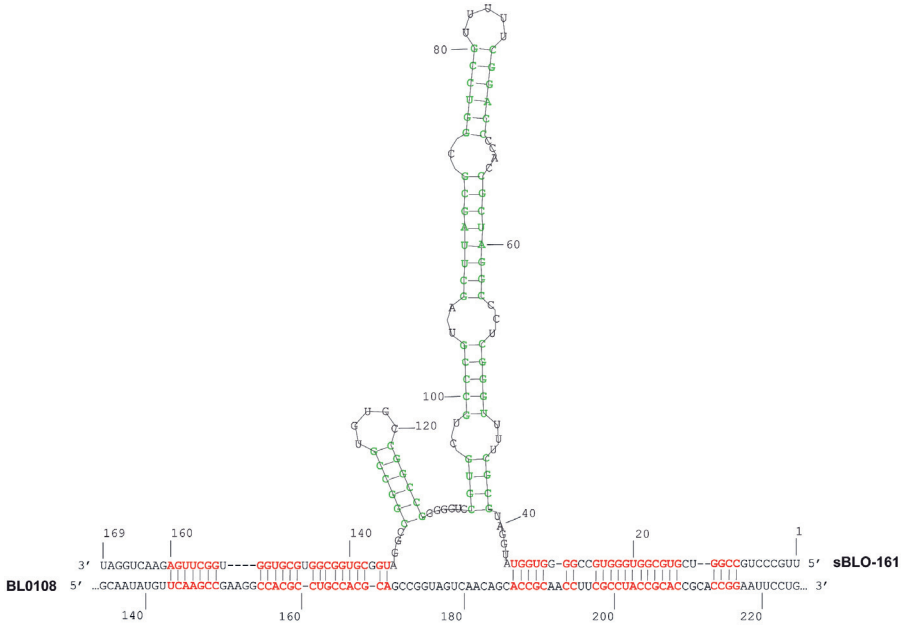


Figure 3: Interaction model of sBLO-161 with the first 240 bp of *BL0108* mRNA sequences. *BL0108* and sBLO-161 bases predicted by intaRNA to bind together are depicted in red. sBLO-161 bases predicted by RNAfold to bind and fold together are depicted in green.

sBLO-161 is predicted to control several key functions in *B. longum* NCC 2705 including central carbon metabolism. RNApredator analysis revealed that sBLO-161 was predicted to target a range of 304 mRNA sequences in *B. longum* NCC 2705 (< -16 kcal/mol), implying that this sRNA may affect a substantial number of genes and play a broad regulatory role in *B. longum* NCC 2705. The gene ontology enrichment analysis performed on the 304 sBLO-161 targeted mRNA supported a role of this sRNA in the regulation of general stress response in NCC 2705, by targeting 6 transcripts involved in nucleotide/base-excision repair (*BL0703; BL0702; BL1025; BL0393; BL1812; BL1207*) (Table S 4). Analysis of the predicted sBLO-161 regulons further strengthened this hypothesis, as it identified as possible targets the transcripts of the general heat-shock protein GrpE (*BL0519*) and its regulator (*hrcA*; *BL0718*), a transfer-messenger RNA binding protein shown to be induced upon heat stress (*smpB*; BL1180) [238], a transporter involved in bile detoxification (*betA; BL1102*), an H+/Na+ antiporter (*BL0942*) that may play a role in salt resistance, the ferredoxin encoding gene (*BL1536*) previously suggested to be part of oxidative stress response [276], and the Whib-like transcriptional regulator WblE (*BL1493*) which has been shown to be upregulated in bifidobacteria under a variety of stress conditions [277]. Finally, the gene ontology analysis revealed the enrichment in various pathways and processes (gene ontologies) that are relevant for the adaptation to glucose or galactose as carbon source for growth. This is illustrated by the relatedness and partial overlap of pathways and processes that were differentially expressed in the comparative transcriptome analysis of glucose and galactose growing cells [273] and the predicted functions targeted by sBLO-161. These include, functions associated with membrane and envelope biogenesis like peptidoglycan synthesis (*BL0561; BL0964; BL1343; BL1344*), exopolysaccharide production (*BL0964; BL1229*) and fatty acid metabolism (*BL0129; BL1536; BL1537*) (Table S 4; Figure 4). In addition, among the predicted transcript targets of sBLO-161 were 18 functions associated with carbohydrate and monosaccharide metabolism, further illustrating the relatedness of enriched functional categories among the sBLO-161 targets and those that were previously shown to be differentially expressed in glucose and galactose grown cells (i.e., associated with the GO-categories: cellular carbohydrate, monosaccharide, and galactose metabolic process; Table S 4; Figure 4).

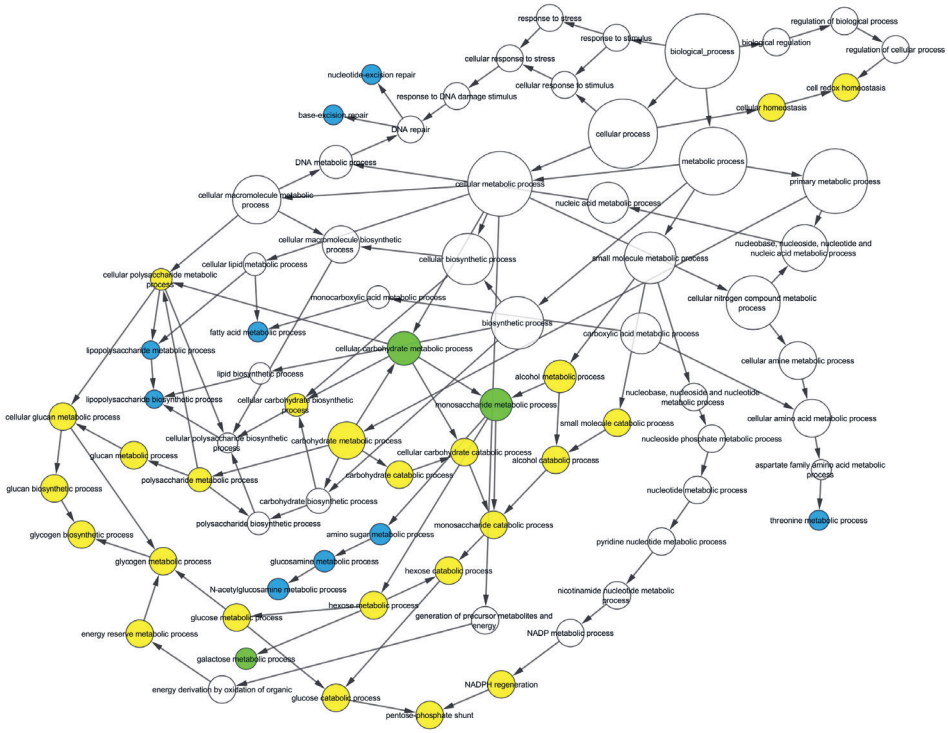


Figure 4: GO functions predicted to be regulated by sBLO-161 (based on regulons with energy of interaction <-16 kcal / mol) as compared to the ones enriched in galactose grown cells. GO found to be enriched in the transcriptome analysis of galactose grown cells [273] are colored in yellow, sBLO-161 enriched GO are colored in blue, and GO found to be enriched in both datasets are colored in green.

Remarkably, RNApredator indicated that no less than 12 transcripts of sugar transport genes [19] were targeted by sBLO-161, which was confirmed by analyzing the predicted interaction structures using intaRNA. Moreover, the latter analysis also revealed that 7 of these 12 transcripts may form RNA duplex structures with sBLO-161 that strongly resembled the structure predicted for the interaction between sBLO-161 and the serpin encoding *BL0108* transcript (Table S 5; Figure 3). The sugar transport associated transcripts targeted by sBLO-161 encode proteins that participate in multiprotein transporter complexes that are annotated to import fructose, ribose and xylose (*BL0035*), raffinose (*BL1523*), galactans (*BL0261*), xylans and arabinans (*BL0424* & *BL0426*), xylosides (*BL1704*), and galactose (*BL0165*). Intriguingly, in many of these cases the expression of genes encoding other constituents of the multiprotein import complex were previously identified as being regulated by differential transcription in glucose or galactose growing cells (Figure 5 A) [273]. These findings suggest that transcriptional control and sBLO-161-mediated post-transcriptional control of protein production may act in concert to fine-tune the expression of multiprotein sugar import systems.

RNApredator predictions of sBLO-161-targeted transcripts included 9 mRNAs involved in the central carbon metabolism of NCC 2705, including three transcripts that were previously shown to be upregulated in galactose-grown compared to glucose-grown cells [273], including functions directly involved in galactose metabolism (*galE1* [*BL1644*]; *galE2* [*BL1671*]), and in glyceraldehyde 4-phopshate conversion (*gap [BL1363]*) (Figure 5 B). In addition, sBLO-161 was predicted to target several other transcripts of genes encoding functions involved in the central carbon metabolism of NCC2705, which were not significantly regulated by glucose versus galactose growth (Figure 5 B). These functions included 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase (*pgm*; *BL1447*), and the xylose/fructose-6-phosphoketolase enzyme (F6PPK*; BL0959*) that plays a pivotal role in the bifid-shunt and is strongly determining carbon flux distribution between C_2 and C_3 fermentation end-products. Notably, using the intaRNA platform we confirmed that 6 of the 9 sBLO-161 targets involved in central carbon metabolism (Table S 5) were predicted to form RNA duplex structures with sBLO-161 that resemble those predicted for sBLO-161 interactions with the carbohydrate-transport (see above) and the serpin encoding transcripts (Figure 3).

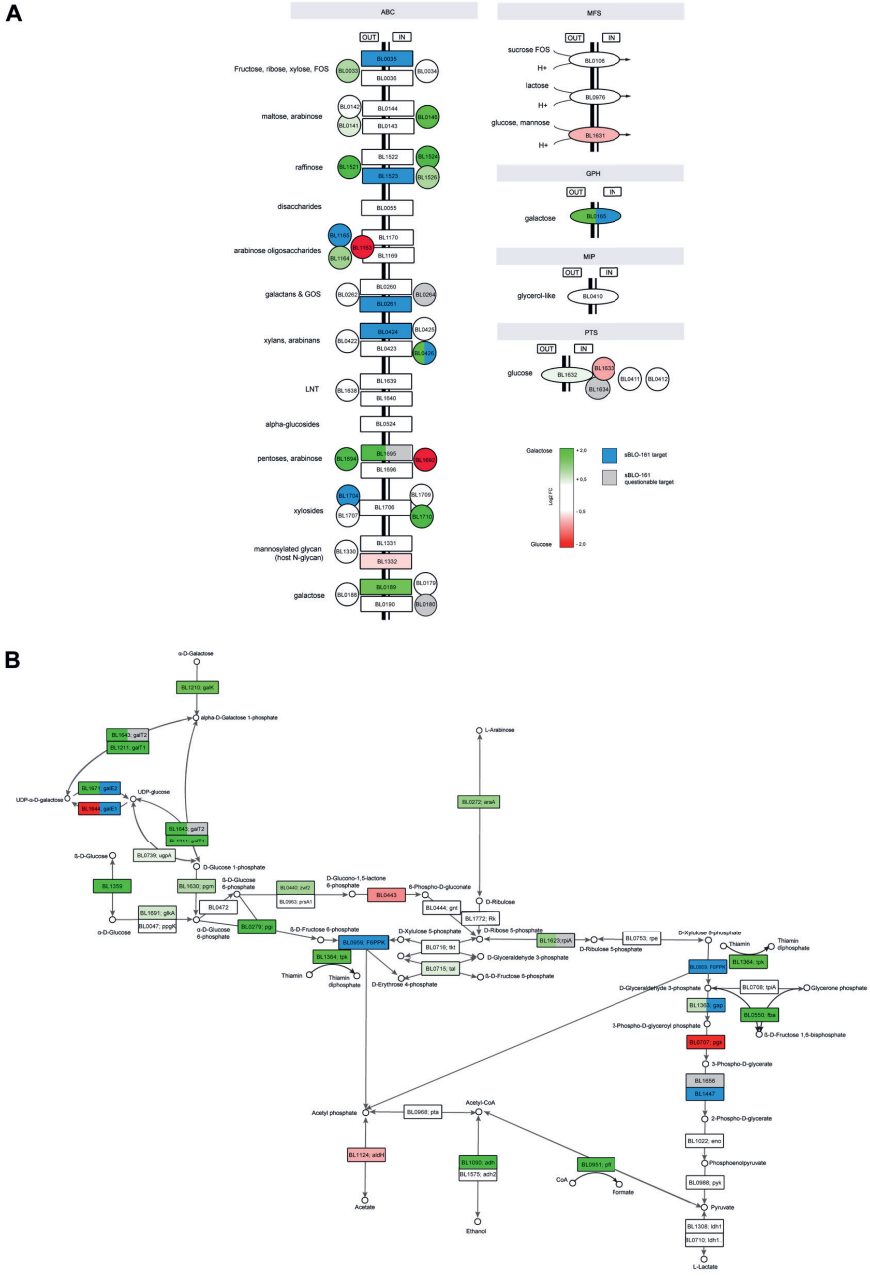


Figure 5: Regulation of sugar import (panel A) and metabolism (panel B) in *B. longum* NCC 2705. Genes previously shown to be induced when the strain was grown on glucose (red) or galactose (green) [273] and transcripts predicted to be targeted by sBLO-161 (blue; energy cut-off < -16 kJ / mol; predictions supported by both RNApredator and intaRNA analysis) are depicted. Questionable sBLO-161 targets based on their interaction profiles proposed by RNApredator and intaRNA are also indicated (grey).

Confirming the translation inhibitory role of sBLO-161 for the F6PPK transcript. The metabolic comparison of *B. longum* NCC 2705 grown on galactose or glucose previously demonstrated that galactose growth was associated with a redistribution of the fermentation end-products, including increased levels of ethanol and formate that predominantly came at the expense of lactate production [273]. Carbon flux distribution within the bifid-shunt likely involves a controlling role of the F6PPK enzyme, which was not regulated at transcriptional level in galactose or glucose grown cells (i.e., 12.7 vs 12.9 log2 counts, respectively [273]) (Figure 6 A), but is predicted to be subjected to differential post-transcriptional control by sBLO-161. To verify the proposed role of sBLO-161 in controlling F6PPK, we measured this enzyme's specific activity in mid-exponentially growing glucose and galactose cultures and confirmed a small, but significative increased F6PPK activity in the latter condition (Figure 6 B). These results corroborate the translation inhibiting role of sBLO-161 when it forms a duplex structure with its target mRNAs. Especially, when the common structural features of this duplex are fulfilled (i.e., mRNA target binding by the 3'- and 5'-ends of sBLO-161 interspaced by an internal stem-loop structure within this sRNA; Figure 3 and Table S 5) it is highly likely that translation of targeted mRNA is inhibited, leading to reduced protein levels in glucose grown cells compared to galactose grown cells.

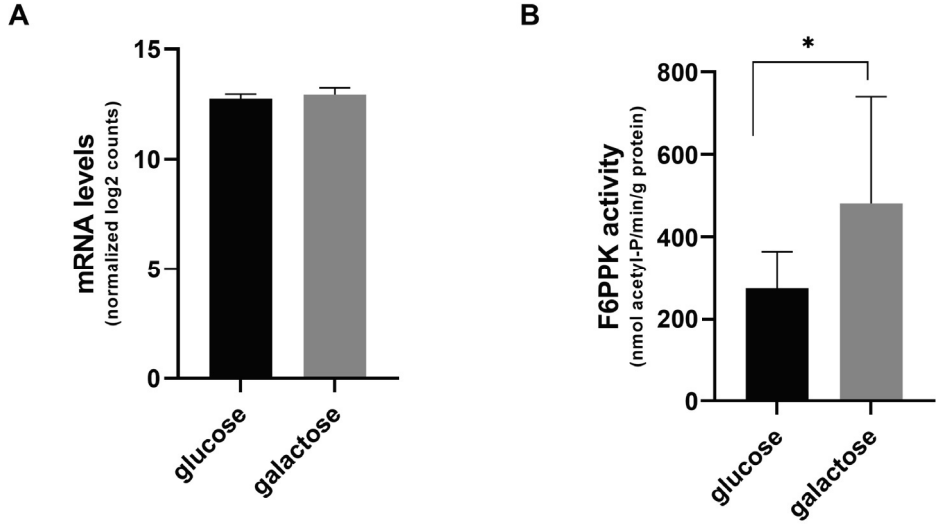


Figure 6: Xylose/fructose-6-phosphoketolase (F6PPK; BL0959) RNA (panel A) and enzyme activity (panel B) levels displayed by mid-exponentially glucose or galactose grown *B. longum* NCC 2705. " p-value<0.05

Evaluating the conservation of the proposed sBLO-161 regulatory role in different B. longum subspecies. BLASTn analysis demonstrated that the *B. longum* NCC 2705 sBLO-161 sRNA sequence is strongly conserved within the available genomes of *B. longum* subsp. *longum* (>98.82% identity), *B. longum* subsp. *suis* and *B. longum* subsp. *suillum* (>93.57% identity), whereas the sequence in *B. longum* subsp. *infantis* appeared to be more distantly related (85% coverage, 87-95% identity). Conservation of the sBLO-161 sequence amongst the *B. longum* spp. was further investigated by alignment of the sBLO-161 homologues encoded by reference strains of the *B. longum* subspecies, confirming the high level of conservation in *B. longum* subsp. *longum* (98.6 % identity), subsp. *suis* and *suillum* (95.1% identity) as well as the lower level of conservation in subsp. *infantis* (89.6 % identity) (Figure 7 A). Notably, sequence divergence in *B. longum* subsp. *infantis* was mostly localized within the 5'-end of the predicted sBLO-161 sequence (Figure 7 B).

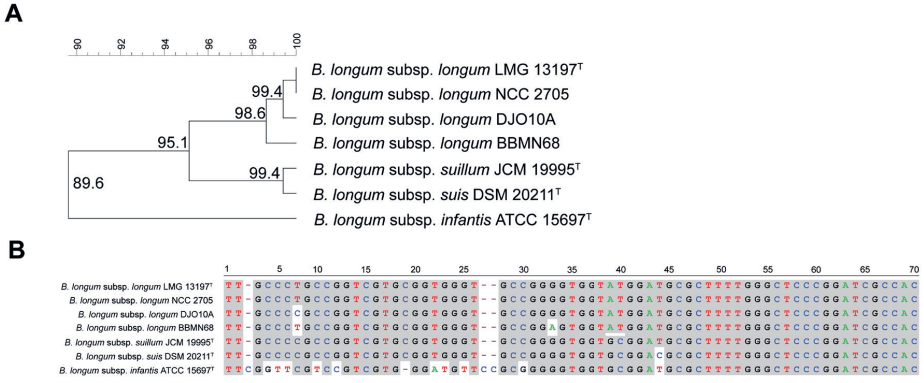


Figure 7: Comparison of the sBLO-161 homologous sequences in *B. longum* subsp. *longum*, *B. longum* subsp. *suis*, *B. longum* subsp. *suillum* and *B. longum* subsp. *infantis* reference strains. Panel A depicts an Unweighted Pair Group Method with Arithmetic mean (UPGMA) tree constructed based on the complete 170 bp sequence alignment of each sBLO-161 homologues. Panel B shows a multiple alignment of the hypervariable region of sBLO-161 (first 70 bp).

As our study suggests that the 5'-end of sBLO-161 plays a critical role in establishing RNA duplexes in a range of the sBLO-161 mRNA targets, the divergence observed in the subsp. *infantis* may affect its role and/or target repertoire in this subspecies compared to subsp. *longum*. Therefore, using the sBLO-161 homologous sequence of *B. longum* subsp. *infantis* ATCC 15697 to predict its targets in the strain's genome using RNApredator, we observed that the sBLO-161 homologue could target 104 transcripts (energy cut-off used <-16.6 kcal/mol). A gene ontology enrichment analysis of the targeted genes identified among others the enrichment of functions related to transposition (5 transposases) and proteolysis (5 proteases). Further analysis of the sBLO-161 targeted transcripts in *B. longum* subsp. *infantis* ATCC 15697 identified 17 genes involved in carbohydrate metabolism, including 9 carbohydrate transport functions and 4 metabolic enzymes, including bifid-shunt associated enzymes (*tkt [BL1096]*, *galE [Blon_2127]*) and enzymes involved in the metabolization of human milk oligosaccharides (*fuconate dehydratase [Blon_2340]*). Of the predicted 104 targets in *B. longum* subsp. *infantis* ATCC 15697 there were 35 targets that have no homologues in *B. longum* NCC 2705, while 27 transcript targets were shared with NCC 2705, including carbohydrate transport (*Blon_0324*, *Blon_0571*, *Blon_1480*) and galactose metabolism (*galE*, *Blon_2127*) functions, but also stress response associated functions like nucleotide/base-excision repair (*Blon_0277*), transfer-messenger RNA binding protein (Blon_2023), Na+/H+ antiporter (*Blon_1705*) or ferredoxin (*Blon_2252*) (Figure 8).

Notably, in *B. longum* subsp. *infantis* ATCC 15697, the sBLO-161 homologue was not predicted to target the transcript of the serpin gene (*Blon_0790*) and its energy of interaction with the F6PPK transcript (*Blon_1722*) appeared relatively low (-9.20 kcal/mol). The lack of targeting to the serpin transcript appeared to be predominantly related to a loss of 5'-end binding capacity (energy of serpin transcript interaction -7.63 kcal/mol) while the 3'-end binding capacity appeared virtually unaffected (-13.74 kJ/mol) (Figure S 3 A, B). Although with a predicted low affinity (i.e., -9.20 kcal/mol) the sBLO-161 homologue in subsp. *infantis* could still form the apparently common duplex structure with the F6PPK transcript (5'end binding -11.84 kcal/mol; 3'end binding -10.32 kcal/mol; with the interspaced stem-loop structure; Figure S 3 panel C, D), which points to the conservation of its role in post-transcriptional regulation of F6PPK expression.

These results overall indicate a relatively well conserved sBLO-161 sequence within *B. longum* subsp. *longum*, *B. longum* subsp. *suis*, and *B. longum* subsp. *suillum.* In *B. longum* subsp. *infantis*, the sBLO-161 homologous sRNA sequence displayed a modified 5'-end sequence which appears to affect its function considerably.

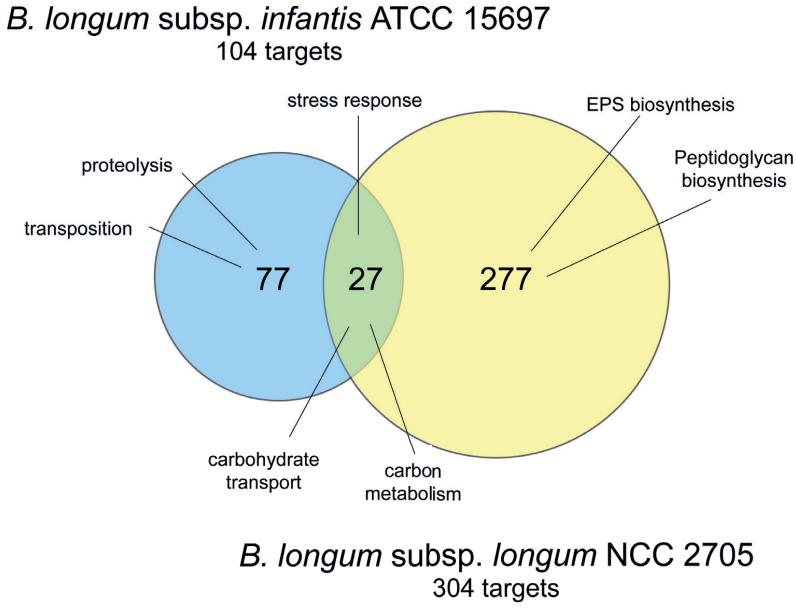


Figure 8: Summary of sBLO-161 regulated functions in *B. longum* subsp. *longum* NCC 2705 and *B. longum* subps. *infantis* ATCC 15697. The number of targets displaying an energy of interaction <-16 kcal /mol predicted using RNApredator are depicted for both strains. Functions identified using gene ontology enrichment analysis in both regulon dataset are also shown.

Conclusions & discussion

This study was inspired by the observation that the strong induction of serpin protein production in galactose grown *B. longum* NCC 2705 compared to its glucose grown counterpart was not explained by an appreciable difference in the serpin encoding mRNA levels. This also appears to contradict the role of AraQ in regulating serpin-gene production by interacting with its promoter region (P_{BL0109}), which we proposed in an earlier study [140], and argued in favor of regulation of serpin production via a post-transcriptional mechanism.

Using paired-end RNAseq data obtained from glucose grown *B. longum* NCC 2705, we detected and quantified a set of 148 candidate sRNA in this strain. Amongst them, sBLO-161 that is located within the intergenic region of the 16S and the 23S rRNA encoding genes, displayed a highly consistent length in all samples (168-169 bp), was significantly overexpressed in glucose grown cells as compared to galactose grown cells, and was predicted to stably bind to the serpin (*BL0108*) mRNA. The predicted RNA-duplex structure formed by the interaction of sBLO-161 and the *BL0108* transcript included the binding of the sRNA 5'- and 3'-ends to the 5'region of the transcript, including the formation of a stem-loop structure within the sRNA sequence. This structure is likely to inhibit effective translation of the *BL0108* transcript and thereby to suppress the production of serpin in glucose grown cells where the expression of the sBLO-161 is substantially higher compared to galactose grown cells. Thereby the sBLO-161-mediated translation inhibition mechanism would explain the observed differential serpin protein production in glucose and galactose grown cultures of *B. longum* NCC 2705 [140].

Strikingly, sBLO-161 in *B. longum* NCC 2705 was predicted to bind more than 300 transcript targets, using a free energy threshold previously employed to predict sRNA targets [274]. This finding suggests that sBLO-161 plays a prominent role in controlling a broad range of transcript translation in this strain, including genes that are annotated to fulfil various cellular function categories, i.e., stress response, synthesis of cell envelope constituents, sugar transport and central carbon metabolism. This global regulatory role may explain why attempts to genetically engineer a stable sBLO-161 overexpressing recombinant strain were unsuccessful. We transformed NCC 2705 with a recombinant plasmid (pSDU17) in which the sBLO-161 sRNA sequence was

fused to a previously described ρ-independent transcriptional terminator sequence [278] and cloned under control of the constitutive P_{gap} promoter [212]. The few transformants obtained displayed a modified colony appearance (i.e., translucid), an altered cell morphology (highly aggregated and irregular Y shape cells) and an impaired capacity to grow in a liquid medium. Subsequent sequencing of the recombinant plasmid extracted from these cells demonstrated that various re-arrangements had taken place, which likely interfered with sBLO-161 expression. We suggest that the proposed global regulatory role of sBLO-161 explains why its overexpression failed as this would lead to pleiotropic effects on cellular physiology of the host strain.

Previous studies have proposed a role of sRNAs in the regulation of stress responses in several bacterial genera [260], including *B. animalis* subsp. *lactis* [271]. Also, sRNAs were shown to control the carbon metabolism in some enteric pathogens [279]. With this study, we propose that a single sRNA (i.e., sBLO-161) is involved in controlling the expression of an unprecedented diversity of gene-function repertoires in *B. longum* NCC 2705, which may include the post-transcriptional control of F6PPK, an enzyme that plays a pivotal role in carbon flux distribution through the “Bifid shunt” in the central carbon metabolism of bifidobacteria. The predicted genome wide regulatory role of sBLO-161 would be in agreement with the pleiotropic physiological consequences associated with the reduced growth rate on galactose compared to the fast growth on glucose [273], conditions that we demonstrated to differentially regulate the sBLO-161 expression level. Moreover, when combining the predicted gene-target repertoire of sBLO-161 with previously analysed differentially expressed genes in glucose and galactose grown cells [273], the complementarity and cooperation of transcriptional and post-transcriptional (i.e., involving sBLO-161) regulation mechanisms in the adaptation to growth on these different carbon sources becomes striking.

The sBLO-161 sRNA is encoded within the 16S-23 rDNA intergenic region, which has been used for species or subspecies delineation in different studies [280,281]. Our analysis established a substantial level of conservation of the sBLO-161 sequence within the *B. longum* species with nevertheless limited subspecies specific sBLO-161 sequence divergence in the 5’-end binding region of the sRNA sequence. In *B. longum* subsp. *longum*, *suis*, and *suillum*, these 5’-end sequence variations are minimal while they are substantial in *B. longum* subsp. *infantis.* We predicted these sequence modifications to be associated with a partial modification of the sBLO-161 orthologue

regulon repertoire in this subspecies. Notably, in *B. longum* subsp. *infantis* the sBLO-161 orthologue was predicted to not regulate serpin production, which is in agreement with the observation that the carbohydrate-mediated regulation of serpin production is not conserved in strains belonging to *B. longum* subsp. *infantis* [140]. Nevertheless, and despite the 5'-end sequence divergence, the *B. longum* subsp. *infantis* orthologue of sBLO-161 is predicted to still exert part of the regulatory role it is proposed to play in *B. longum* subsp. *longum* NCC 2705, including its post-transcriptional regulation role of several stress responses, carbohydrate transport and metabolism functions.

Overall, this study established the existence of an important regulatory sRNA in a region of the genome that is conserved at *B. longum* subspecies level. Our data point to a plausible more general role of the sBLO-161 sRNA in the regulation of a diverse range of microbial functions and hence highlight the interest to pursue studies aiming at deciphering the role of sBLO-161 homologues in other genera and species.

Methods

***B. longum* NCC 2705 cultivation.** *B. longum* NCC 2705 was obtained from the Nestle Culture Collection (NCC) (Nestlé Research, Lausanne, Switzerland). The strain was grown in a previously described homemade variant of the de Man, de Rogosa, Sharpe medium (MRS) [140] . The basic medium containing no carbohydrates (MRS-C) and was supplemented with 1% galactose or 1% glucose (Sigma-Aldrich Chemie GmbH, Buchs, Switzerland). Carbohydrate stock solutions were all prepared at 100 g/L in water and filtered sterilized using a 0.22 µm filter. Cultures were inoculated with 2% of an overnight pre-culture. Glucose and galactose-based fermentations were performed in triplicate (6 cultures in total) in two liter DasGip bioreactors (Eppendorf AG, Hamburg, Germany) using a 500 ml working volume. Anaerobiosis was achieved by CO2 flow of the headspace, agitation was maintained at 600 rpm and temperature was set to 37°C. pH was monitored inline during the entire fermentation. Cells were collected in mid-exponential phase and harvested by centrifugation (3000 g, 10 min at 4°C). Bacterial pellets were stored at -80°C until further analysis.

Serpin measurement. Anti-serpin sandwich ELISA was performed as described by Duboux et al [140]. Briefly, 96-well plates (Nunc MaxiSorp [Thermo Fisher Scientific AG, Basel, Switzerland]) were coated with anti-serpin antibodies. After washing and blocking of the wells, samples were applied and a titration range of purified recombinant serpin [61] was used to generate a standard curve. Bound serpin protein was detected using biotinylated anti-serpin antibodies. Detection was performed by addition of streptavidin-labelled horse-radish peroxidase (HRP Pierce, Thermo Fisher Scientific AG) enzyme conjugate, followed by washing to remove background reactivity. Chromogenic detection was performed by addition of HRP substrate (1-step Ultra TMB ELISA, ThermoFisher Scientific AG) followed by 15 minutes incubation. The reaction was terminated by addition of 100 µl of 2M sulfuric acid followed by optical density measurement at 450 nm in a Varioskan spectrophotometer (Thermo Fisher Scientific AG).

RNA preparation and sequencing. RNA preparation and sequencing was performed as described previously [273]. Briefly, RNA was isolated from mid-exponentially grown cells using the miRNeasy kit (QIAGEN AG, Hombrechtikon, Switzerland). Library preparations were prepared with the Prokaryotic AnyDeplete kit (Protocol M01502 v1,

NuGEN Technologies Inc., Redwood City, USA) starting from 250 ng of RNA for each sample. Quality of the libraries was checked prior sequencing on an Illumina HiSeq 2500 with Rapid V2 chemistry.

sRNA detection and quantification in glucose or galactose grown *B. longum* NCC 2705. Detection and quantification of small transcripts from the paired-end RNA sequences (on all six samples) obtained above was performed using the APERO software [267]. For this purpose, the APERO package was installed following the github instructions (https://github.com/Simon-Leonard/APERO) and tested using the test data and R commands provided by the developers (https://github.com/Simon-Leonard/APERO/tree/master/example). Transcript size and read-quantity detected by APERO in the genomic regions encompassing sBLO-161, sBLO-1969 and sBLO-1578 were evaluated in all six runs. Transcripts copies detected by APERO for those three candidate sRNA in both the glucose and galactose conditions were subjected to a two-tailed Welch t-test in GraphPad Prism (v8.1.1, GraphPad Software Inc., San Diego, USA).

sRNA targets, RNA-RNA interaction and RNA folding predictions. Initial selection of sRNA targeting *BL0108* was performed using RNAseq derived transcripts detected in a glucose grown sample. We first discarded small transcripts located inside protein-coding sequences ("I", Figure S 1) as those might correspond to mRNA degradation products. Remaining transcripts were further filtered based on size (between 50-350 bp) and frequency of detection (>50) resulting in a final list of 198 transcripts (Table S). Sequences of these transcripts were retrieved from the publicly available genome of *B. longum* NCC 2705 [16], using the 5' and 3' ends detected by APERO. These sequences were then used as input in the RNApredator web-interface [269] to predict the mRNA they target in the genome of *B. longum* NCC 2705. From this initial step, all sRNA displaying a stable binding interaction with *BL0108* were selected using an energy of interaction threshold of -16 kcal / mol, which is a threshold previously proposed for the detection of sRNA targets [274]. sRNA obtained at this stage were named according the nomenclature system proposed by Li and colleagues [275]: an initial 's' which stands for small RNA, followed by a three-letter genome ID used in the KEGG database (in the case of NCC 2705: BLO), and a number that indicates its genomic location (i.e. the position in kbp +1).

Detailed RNA-RNA interactions were computed using the available online prediction software IntaRNA [270]. The sRNA sequences and their predicted targeted sequences were used as input. RNA-RNA interactions were predicted using the standard software setting. Folding of the central sequence of both sBLO-161 (positions 55 to 149) and sBLO-1969 (positions 31 to 104) was predicted using the online available tool RNAfold, hosted by the University of Vienna (http://rna.tbi.univie.ac.at/). Finally, a model of interaction between the first 250bp of *BL0108* and the sBLO-161 was constructed in Adobe Illustrator (v24.3; Adobe Inc., San José, USA) based on both IntaRNA and RNAfold predictions.

sRNA sequence comparisons in publicly available *B. longum* genomes. The 168-9 bp long sequence of sBLO-161 was first compared to all available *B. longum* reference genomes using the BLASTN platform hosted at the National Center for Biotechnology Information (NCBI; https://blast.ncbi.nlm.nih.gov/). Homologous sBLO-161 sequences in representative strains (*B. longum* subsp. *longum* LMG 13197^T, DJO10A, BBMN68; *B. longum* subsp. *suillum* JCM 1995^T; *B. longum* subsp. *suis* DSM 20211^T; *B. longum* subsp. *infantis* ATCC 15697^T) were retrieved manually, aligned and compared using a UPGMA based tree in the BioNumerics software (v8.0, bioMérieux SA, Marcy l'Etoile, France).

Fructose-6-phosphate phosphoketolase (F6PPK) activity measurement. Briefly, exponentially grown glucose or galactose NCC 2705 cells were harvested by centrifugation (6,000 g, 5 min, 4 °C) and washed twice with ice cold PBS. The pellet was resuspended in 800 µL ice-cold PBS supplemented with 1X SigmaFast Protease Inhibitor (Sigma-Aldrich Chemie GmbH) and cell lysates were obtained by beating using FastPrep-24 machine (MP Biomedicals SARL; 3 cycles of 1 minute at 4 m/s). The solubilized protein fraction was collected by centrifugation (12,000 g, 10 min, 4 °C) and the total protein content was quantified using the Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific AG) following manufacturer's instructions. The F6PPK enzyme activity of the lysate was measured according to a modified version of the assay described by Orban and Patterson [282]. Twenty-five µL of a solution containing sodium fluoride and sodium iodoacetate (3 and 5 mg/mL, respectively) and 25 µL of D-fructose-6-phosphate solution (80 mg/mL) (Sigma-Aldrich Chemie GmbH) were added to 500 µL of lysate. The solution was vortexed and subsequently incubated at 37 °C for 30 minutes. To release inorganic phosphate from the enzymatically produced acetyl-phosphate, 300 µL of 13% (w/v) hydroxylamine-HCl (Sigma-Aldrich Chemie GmbH) were added, and the solution incubated at room temperature for 10 minutes. After incubation, 200 µL each of 15% (w/v) trichloroacetic acid and 4 M HCl were added to precipitate the proteins. Next, 200 µl of 5% (w/v) $FeCl_3 \cdot 6H2O$ (Sigma-Aldrich Chemie GmbH) in 0.1 M HCl were added to catalyze the formation of $FePO_4$. Finally, $FePO_4$ quantification was done spectrophotometrically at 505 nm (Varisokan; Thermo Fisher Scientific AG) by using a set of acetyl phosphate standards (0-62.5 mg/ml). One unit of F6PPK activity was defined as the amount of acetyl phosphate (in nmol) released per mg of total protein in 1 minute.

Supplementary data

Table S 1: List of the 198 small RNA transcripts detected by APERO in a glucose mid-exponentially grown culture of *B. longum* NCC 2705. Transcripts were filtered based on length (50-350 bp), and frequency of detection (>50). Definition of sRNA classes can be found in Figure S 1.

5' end position	length (bp)	strand	frequency	class	genetic context
20901	119	-	455	P	
41898	267	+	375	5U	upstream BL0033
55529,5	182,5	-	61	P	upstream BL0044
57495	266	-	104	I	whithin BL0046
61463,5	184,5	+	134	Ai	antisense to BL0050
63106	240	-	114	I	whithin BL0050
113552,5	213,5	+	119	5U	
115486	204	+	597	5U	upstream BL0096
140960,5	147,5	+	59	Div	diverging with BL0115
151892	264	+	229	I	
160871	169	+	434	O	after Blr01, before Blr02
160946,5	93,5	+	934	O	after BL0130, before BL1293
161777	246	-	1202	O	BL1292, before BL0126
162420	251	-	639	O	after BL1292, before BL0126
163486	179	-	79	O	after BL1292, before BL0126
174586,5	317,5	-	51	I	whithin BL1301
181954	169	-	108	I	whithin BL1305
192031	288	-	57	P	upstream BL1314
202504	170	+	171	I	whithin BL1323
206570	284	+	825	O	after BL1324, before BL1326
208163	277	-	725	P	upstream BL1325
238246	236	-	76	P	upstream BL1342
253156	188	+	863	I	Whithin BL1354
269940	234	+	57	5U	upstream BL1368
269967	207	+	106	5U	upstream BL1368
313150	180	+	348	5U	
314192	194	+	1848	I	whithin BL1405
328534,5	144,5	-	174	Ai	antisense to BL1417
328696,5	252,5	+	1306	5U	upstream BL1418
377132	153	+	81	5U	upstream BL1453
383127	256	-	81	O	after BL1459, before BL1446
422171	147	-	756	I	whithin BL1488
444569	238	-	468	P	upstream BL1499
449344	271	+	5494	5U	upstream BL1503
468652	289	+	50	5U	upstream BL0697
485663	188	+	107	I	whithin BL0707
492952	188	+	113	3A	
497888	266	-	4185	P	upstream BL0716
501566	124	+	77	I	
503367	167	-	152	I	whithin BL0722
517875,5	269,5	-	109	P	upstream BL0736
527917	153	+	220	5U	upstream BL0743
550009	326	-	53	I	whithin BL0768
555959	131	+	334	5U	upstream BL0778
559251,5	117,5	+	179	5U	upstream BL0781
561320	281	+	283	5U	upstream BL0783
565181	243	+	113	I	whithin BL0786
574336	240	-	817	P	upstream BL0794
582040	189	-	69	P	upstream BL0798
584003	173	-	136	P	upstream BL0800
635368	118	-	182	P	upstream BL0852
636225	207	-	761	P	
668142	332	-	61	I	whithin BL0880
693167	271	-	76	I	whithin BL0901
697441	150	-	107	P	upstream BL0905
733623	234	-	1069	P	upstream BL0932
747108	265	-	79	P	upstream BL0942
759534,5	265,5	-	738	I	within BL0951
767800	264	+	425	Ai	antisense to BL0959
770039,5	292,5	-	11571	P	upstream BL0959
772246	148	-	193	O	after BL0970, before BL0959
800427,5	156,5	-	542	I	Whithin BL0984
806886	348	-	3787	I	Whithin BL0988
812613	239	-	3410	P	upstream BL0992
860066	324	-	88	I	Whithin BL1034

872474	327	-	123	P	upstream BL1044
872652	315	+	67	5U	
882398	272	-	333	P	upstream BL1053
889940	165	-	125	P	upstream BL1060
925771,5	165,5	+	265	5U	
935113	324	-	3528	P	
937273	318	+	1577	5U	upstream BL1103
939591,5	230,5	-	63	I	Whithin BL1104
955136,5	188,5	-	243	P	upstream BL1114
958921	192	+	1650	5U	upstream BL1118
961695	215	+	65	O	after BL1120, before BL1121
981351	232	+	83	5U	upstream BL1136
986014	258	+	1262	I	Whithin BL1139
986060	213	+	3876	I	Whithin BL1139
987564	98	-	186	P	upstream BL1140
1001146	239	-	302	P	upstream BL1152
1015521,5	141,5	-	58	P	upstream BL1163
1055771	253	-	82	P	
1055807,5	289,5	-	439	P	
1095024,5	283,5	-	3111	I	Whithin BL1212
1111444	195	+	59	I	
1111988	147	+	1768	5U	upstream BL1227
1135589,5	273,5	-	146	I	Whithin BL0229
1156391	257	-	105	P	upstream BL0248
1189194,5	250,5	+	210	I	
1194343,5	315,5	-	65	P	upstream BL0273
1199306,5	263,5	-	149	P	upstream BL0277
1207925	229	-	586	P	upstream BL0284
1214451	224	-	72	O	after BL0299, before BL0289
1216914	173	+	217	5U	upstream BL0294
1220791	289	+	60	O	after BL0297, before BL0298
1220860,5	219,5	+	312	O	after BL0297, before BL0298
1227568	116	-	65	I	
1251127	242	-	708	P	upstream BL0322
1260538,5	149,5	-	439	P	upstream BL0329
1269662	112	+	335	I	whithin BL0339
1270280	306	+	86	3U	
1315580,5	142,5	+	741	I	
1323606	333	+	163	I	whithin BL0395
1331560	306	-	272	P	upstream BL0398
1333283,5	267,5	-	1042	I	Whithin BL0399
1348249	285	-	58	I	Whithin BL0411
1371073	160	+	55	I	Whithin BL0426
1374222	307	-	85	I	Whithin BL0428
1395971	148	-	68	P	upstream BL0445
1424895	90	-	170	O	after BL0469, before BL0464
1424995	189	-	135	O	after BL0469, before BL0464
1427724,5	221,5	+	59	5U	
1434231,5	284,5	+	559	I	
1465203	133	-	205	I	Whithin BL0500
1500654	135	-	196	P	upstream BL0532
1551726	164	-	3943	P	upstream BL0576
1571617,5	95,5	+	223	5U	
1577851	195	-	80	I	Whithin BL0599
1577973	264	+	127	5U	
1590578	241	+	65	I	Whithin BL0608
1599631	202	+	131	5U	upstream BL0617
1601590	299	+	62	5U	upstream BL0619
1606171,5	176,5	+	50	O	after BL0622, before BL0629
1610124	221	-	57	P	upstream BL0628
1615624	282	-	106	P	upstream BL0633
1620570	176	-	533	P	upstream BL0635
1623655,5	175,5	+	59	Ai	
1640828	202	-	1853	O	after BL0655, before BL0653
1655174	222	+	253	I	
1705739	207	+	170	O	after BL1527, before BL0140
1705925	279	+	77	O	after BL1527, before BL0140
1706202	278	-	77	O	after BL0137, before BL1528
1707100	144	-	139	O	after BL0137, before BL1528
1707273	317	-	74	O	after BL0137, before BL1528
1707530	168	+	1063	O	after BL1527, before BL0140
1707606,5	91,5	+	404	O	after BL1527, before BL0140
1708435	233	-	112	O	after BL0137, before BL1528
1708807,5	271,5	+	1837	O	after BL1527, before BL0140
1708824	255	+	1901	O	after BL1527, before BL0140

1708885	194	+	2207	O	after BL1527, before BL0140
1709971	189	+	32928	O	after BL1527, before BL0140
1709991	183	+	4512	O	after BL1527, before BL0140
1714376	237	+	2427	O	after BL1527, before BL0140
1714452	338	+	68	O	after BL1527, before BL0140
1714555	235	+	561	O	after BL1527, before BL0140
1714639	210	+	1099	O	after BL1527, before BL0140
1714735	243	+	574	O	after BL1527, before BL0140
1714982	273	+	61	O	after BL1527, before BL0140
1715036	219	+	181	O	after BL1527, before BL0140
1715053	202	+	100	O	after BL1527, before BL0140
1715080	175	+	78	O	after BL1527, before BL0140
1715095,5	159,5	+	267	O	after BL1527, before BL0140
1715146	124	+	7020	O	after BL1527, before BL0140
1715190	202	+	612	O	after BL1527, before BL0140
1716139	196	+	233	O	after BL1527, before BL0140
1716182,5	178,5	+	97	O	after BL1527, before BL0140
1716731,5	300,5	+	81	O	after BL1527, before BL0140
1731177	241	-	64	P	upstream BL0142
1761394	111	-	175	O	after BL0159, before BL0153
1774813	121	+	71	3A	antisense to BL0171
1785559,5	177,5	+	1284	P	upstream BL0179
1807756	173	+	357	I	Whithin BL0193
1815722	239	+	123	5U	
1835217,5	323,5	+	677	I	
1865927	175	-	907	I	Whithin BL1258
1897407	294	+	179	5U	upstream BL1281
1908739	205	+	871	O	after BL1291, before BL1529
1908925	279	+	318	O	afer BL1291, before BL1529
1910530	168	+	455	O	after BL1291, before BL1529
1910606,5	91,5	+	4271	O	after BL1291, before BL1529
1911436	227	-	50	O	after BL1531, before BL1289
1912909,5	304,5	+	59	O	after BL1291, before BL1529
1912963	191	+	1587	O	after BL1291, before BL1529
1913006	196	+	1264	O	after BL1291, before BL1529
1913451	185	+	85	O	after BL1291, before BL1529
1913636,5	221,5	+	164	O	after BL1291, before BL1529
1913680	226	+	71	O	after BL1291, before BL1529
1913703	205	+	137	O	after BL1291, before BL1529
1921185	174	-	269	P	upstream BL1534
1956648	274	-	869	P	
1966594	101	+	198	I	
1968821,5	146,5	+	4772	5U	upstream BL1558
1971379	95	+	87	Div	diverging with BL1559
1987364,5	305,5	-	65	Ai	antisense to BL1572
1987783	345	-	58	I	Whithin BL1573
2029483	149	+	210	5U	
2059756	269	+	342	I	
2077191	161	+	385	I	Whithin BL1661
2106889,5	272,5	+	612	5U	upstream BL1686
2127341	271	-	271	O	after BL1709, before BL1703
2130794	109	+	632	5U	upstream BL1711
2159364,5	162,5	+	61	3A	
2180637,5	268,5	+	328	I	Whithin BL1753
2181508	216	+	88	5U	upstream BL1754
2238641	338	+	239	I	Whithin BL1796
2241694	277	-	330	P	upstream BL1798
2255781	339	+	198	I	

Table S 2: List of the 5 putative sRNA predicted to bind the *BL0108* transcript as determined by RNApredator (<-16 kcal/mol). The average size (± standard deviation) of each sRNA detected by APERO in each the 6 biological replicates is depicted. The respective sRNA sequences detected by APERO and used as input for the RNApredator interaction predictions are included.

sRNA name	Detected 5' end position	Detected 3' end position	class	length (bp)	strand	genomic context	mRNA binding positions	sRNA binding positions	energy of interaction (hybridization; rRNA unfolding; mRNA unfolding) (kcal/mol)
sBLO-161	160870-1	161040	O	168-9	+	In between Blr01 (16S) & Blr07 (23S)	157-184	138-162	-21,26 (-44,5 ; 10,4 ; 12,84)
sBLO-1969	1968821	1968968-9	5U	147-148	+	Overlaps start of groES (BL1558)	316-330	15-30	-16,63 (-32,1 ; 8,17 ; 7,3)
sBLO-1578	1577901-74	1578236-39	5U	263-336	+	Overlaps start of trpS (BL0600)	854-880	141-168	-16,95 (-45,7 ; 15,01 ; 13,74)
sBLO-1560	1599629-1599632	1599833-160050	5U	201-419	+	Overlaps start of BL0617	316-337	42-64	-16,97 (-37 ; 11,57 ; 8,46)
sBLO-1611	1610121-26	1609326-920	P	201-799.5	-	Overlaps start of BL0628	363-372	206-215	-17,55 (-24,7 ;3,69 ; 3,46)

Table S 3: Summary of the detailed 5' and 3' end interactions of sBLO-161 within the first 240 bp of *BL0108*, as predicted using intaRNA.

sRNA sequence region	sBLO-161 binding site	BL0108 binding site	energy of interaction (hybridization; rRNA unfolding; mRNA unfolding) (kcal/mol)
5' END	09-36	187-217	-15.96 (-39,3; 7,48; 15.83)
3' END	136-161	143-170	-10.17 (-37.25; 9,99; 17,09)

Table S 4: Gene Ontology (GO) enrichment analysis result obtained with the BINGO plugin of Cytoscape, using the 304 sBLO-161 targeted genes of *B. longum* NCC 2705 and gene ontologies retrieved from UNIPPROT.

GO-ID	GO description	p-value	n° of genes associated to GO	genes in test set
6041	glucosamine metabolic process	4,97E-03	4	BL0964|BL1344|BL1343|BL0561
6044	N-acetylglucosamine metabolic process	4,97E-03	4	BL0964|BL1344|BL1343|BL0561
6040	amino sugar metabolic process	1,28E-02	4	BL0964|BL1344|BL1343|BL0561
6289	nucleotide-excision repair	2,19E-02	3	BL0703|BL0702|BL1025
6566	threonine metabolic process	2,19E-02	3	BL1527|BL1275|BL1274
6012	galactose metabolic process	2,55E-02	4	BL1168|BL1644|BL1643|BL1671
8653	lipopolysaccharide metabolic process	3,45E-02	2	BL0964|BL1229
9103	lipopolysaccharide biosynthetic process	3,45E-02	2	BL0964|BL1229
44262	cellular carbohydrate metabolic process	3,48E-02	18	BL0181|BL1168|BL1344|BL1644|BL1643|BL1363|BL1671|BL1343|BL0561|BL1573|BL1704|BL0826|BL0964|BL1656|BL1623|BL0911|BL1229|BL1328
5996	monosaccharide metabolic process	4,67E-02	16	BL0181|BL1168|BL1344|BL1644|BL1643|BL1363|BL1671|BL1343|BL0561|BL1573|BL1704|BL0964|BL1656|BL1623|BL0911|BL1328
6284	base-excision repair	4,74E-02	3	BL0393|BL1812|BL1207
6631	fatty acid metabolic process	4,74E-02	3	BL1536|BL1537|BL0129

Table S 5: Summary of the sBLO-161 IntaRNA binding profiles to *B. longum* NCC 2705 mRNAs involved in the carbon transport (A) and metabolism (B) of the strain.

GENE ID	UNIPROT annotation	sBLO-161 5' end binding position (energy of interaction in kcal/mol)	sBLO-161 3' end binding position (energy of interaction in kcal/mol)	Comment
BL0035	Fructose import permease protein FruF	28-57 (-19.11)	155-168 (-11.02)	
BL1523	Sugar permease of ABC transporter system	28-57 (-18.89)	151-170 (-13.97)	
BL1165	Probable solute binding protein of ABC transporter system for sugars	9-36 (-13.84)	126-151 (-15.24)	
BL0261	Sugar transport system permease protein	20-46 (-9.98)	137-170 (-12.16)	
BL0264	ATP binding protein of ABC transporter	7-47 (-22.34)	77-170 (-19.87)	very long 3' end binding (no loop formation)
BL0426	LacI-type transcriptional regulator	8-54 (-18.86)	137-151 (-11.43)	
BL1695	ATP binding protein of ABC transporter for pentoses	30-36 (-12.24)	152-163 (-9.83)	short bindings with middle binding site (90-114; -12.62)
BL1704	Xylose isomerase (EC 5.3.1.5)	17-58 (-16.27)	142-170 (-15.38)	
BL0180	Probable permease protein of ABC transporter system	30-52 (-12.98)	142-161 (-14.11)	middle binding site (77-97; -13.24)
BL0165	Possible symporter	2-23 (-16.87)	154-168 (-17.2)	
BL1634	DAGKc domain-containing protein	22-68 (-10.92)	150-158 (-5.65)	no real 3' end binding; binding in the middle (83-107; -9.79)

GENE ID	UNIPROT annotation	sBLO-161 5' end binding position (energy of interaction in kcal/mol)	sBLO-161 3' end binding position (energy of interaction in kcal/mol)	Comment
BL1643	Galactose-1-phosphate uridylyltransferase (Gal-1-P uridylyltransferase) (EC 2.7.7.12) (UDP-glucose--hexose-1-phosphate uridylyltransferase)	28-57 (-19,74)	86-168 (-17,36)	very long 3' end binding
BL1671	UDP-glucose 4-epimerase	8-35 (-24,27)	11-169 (-15,3)	
BL1623	Ribose-5-phosphate isomerase A (EC 5.3.1.6) (Phosphoriboisomerase A) (PRI)	30-48 (-12,28)	152-166 (-8,01)	low 3' end & middle binding (115-126; -10.64)
BL0959	Xylulose-5-phosphate/fructose-6-phosphate phosphoketolase	20-42 (-10,62)	150-168 (-11,25)	
BL1363	Glyceraldehyde 3-phosphate dehydrogenase C	22-36 (-20,03)	150-169 (-12,77)	
BL1656	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase (BPG-dependent PGAM) (PGAM) (Phosphoglyceromutase) (dPGM) (EC 5.4.2.11)	1-21 (-13,64)	133-168 (-18,11)	no real clear binding regions (overall low energy over the whole sequence)
BL1447	Possible phosphoglycerate mutase	8-46 (-15,65)	114-168 (-14,84)	

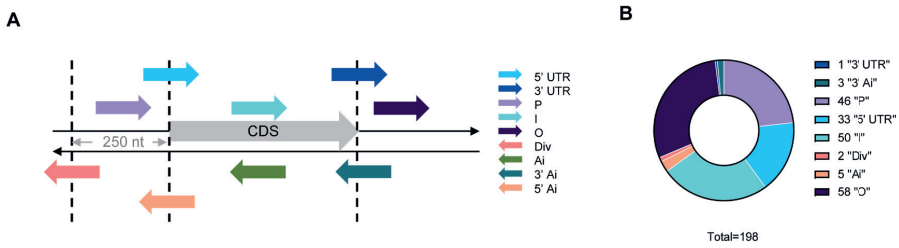


Figure S 1: Small RNA detection results in mid-exponentially glucose grown *B. longum* NCC 2705. Panel A represent the different classes of sRNA predicted by APERO. This figure is adapted from Leonard et al [267]. Panel B summarizes the distribution of hits according to their class, filtered based on size (50-500 bp) and detection frequency (>50).

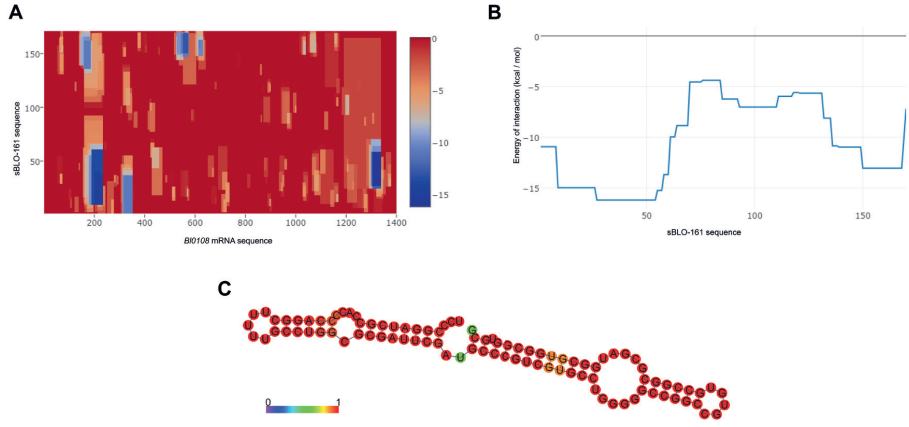


Figure S 2: Predicted binding & folding profiles sBLO-161 to *BL0108* mRNA. Panels A represents the IntaRNA predicted interactions between the target mRNA (BL0108) represented on the X axis, and the sRNA sequences, represented on the Y axis. Interactions with a low energy level are depicted in blue. Panel B represents the binding profiles of the sRNA to the BL0108 mRNA as predicted by intaRNA, with on the X axis the sRNA sequence and on the Y axis the minimal energy corresponding to the interaction with the BL0108 mRNA. Folding of the sequences encompassed between the predicted binding sites was predicted using RNAfold and are represented in panel C. Base-pairing affinities are depicted in different colors, from blue (low) to red (high).

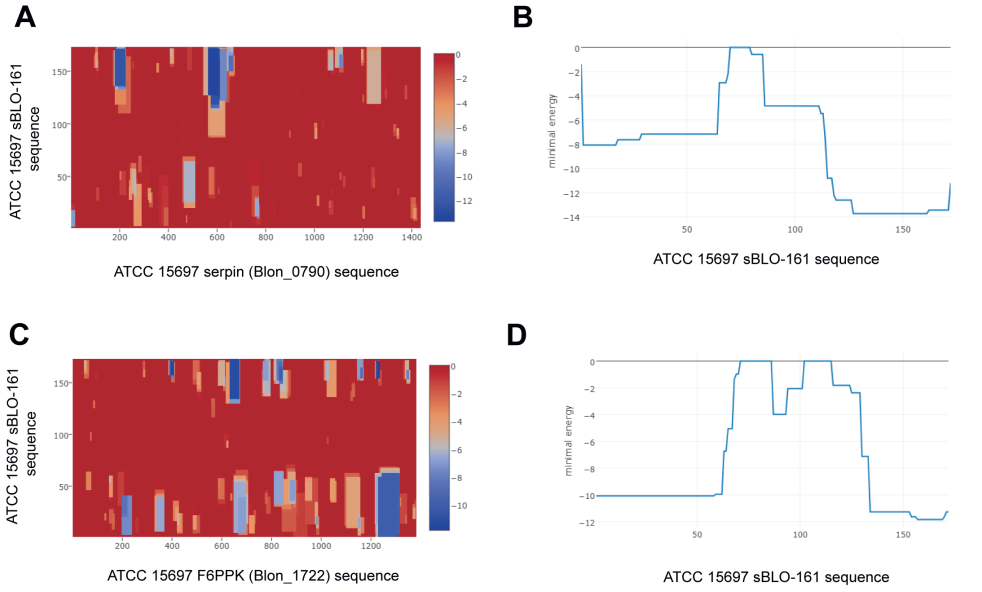
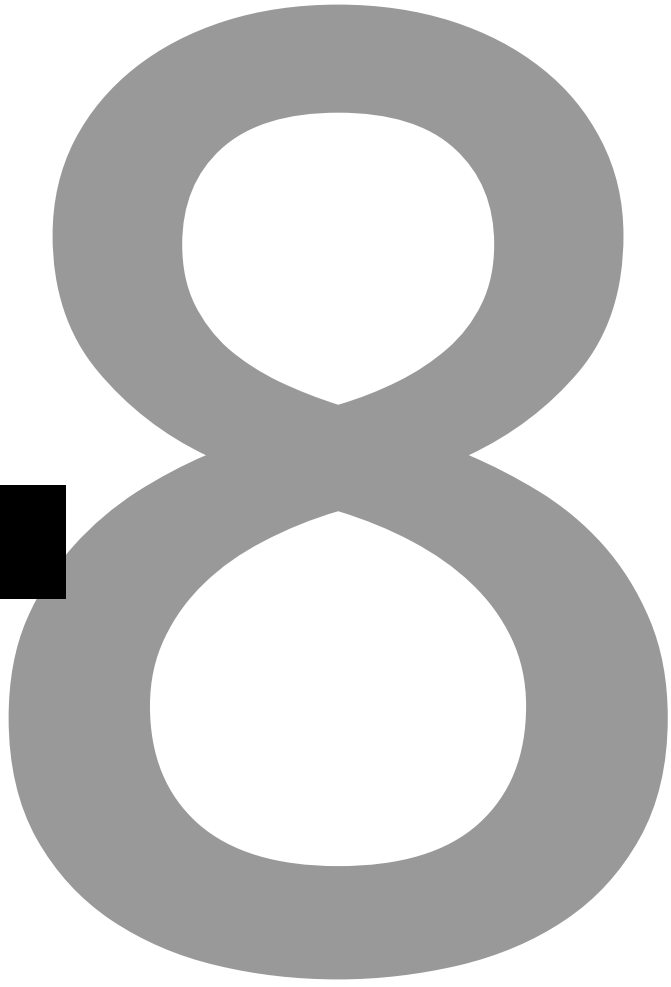


Figure S 3: Predicted interactions of the *B. longum* subsp. *infantis* ATCC 15697 sBLO-161 homologue sequence with the F6PPK (Blon_01722) (Panels A & B) and the serpin (Blon_0790) (Panels C & D) encoding mRNAs. Panels A & C, represent the IntaRNA predicted interactions between the target mRNA represented on the X axis and the sRNA sequences represented on the Y axis. Interactions requiring a low level of energy are depicted in blue. Panels B & D represent the binding profiles of the sRNA to the Blon_1722 and Blon_0790 mRNA as predicted by intaRNA, with on the X axis the sRNA sequence and on the Y axis the minimal energy corresponding to the interaction with the target mRNA.

CHAPTER 8



General discussion

S. Duboux [1,2]

[1] Nestlé Research, Route du Jorat 57, CH 1000 Lausanne 26, Switzerland
[2] Host-Microbe Interactomics Group, Wageningen University & Research, De Elst 1, 6708WD Wageningen, The Netherlands

The work presented in this thesis aims at understanding if, and how, production conditions of *Bifidobacterium* probiotics influence their physiology and bioactivity. Throughout the thesis, we based our evaluations on the expression of bifidobacterial niche factors and effector molecules previously demonstrated to drive the *in situ* establishment or beneficial activities of different strains belonging to this genus. After an initial assessment of the current knowledge in this scientific area (chapters 1 & 2), we used *B. longum* NCC 2705 and the serine protease inhibitor (serpin), a key effector molecule it produces, as a model throughout our different experiments (chapters 3,4,7). In this section, we will summarize and discuss the carbohydrate-controlled mechanism by which serpin production is regulated in *B. longum* strains. Importantly, our different studies demonstrate the existence of a carbohydrate mediated catabolite repression mechanism that mediates pleiotropic physiological effects in NCC 2705. In this chapter we will integrate our different findings to highlight how catabolite repression regulates the carbohydrate metabolism of that strain, including the involvement of a purported global regulatory sRNA we discovered and named sBLO-161. We will further discuss how the production of this important sRNA may be regulated. Finally, we will propose how the data generated in this thesis could be leveraged to establish serpin-dependent bifidobacterial probiotic concepts that deliver health beneficial effects in a predictable and consistent manner.

The regulatory mechanisms controlling *B. longum* serpin production

Prior to our studies, bifidobacterial serpin regulation knowledge was limited to *B. breve*, were the transcription of the serpin encoding gene is triggered by exposure of the cells to different proteases, controlled by a two-component regulatory system (named *serRK*) that is absent from *B. longum* strains [137,139]. In line with the reported absence of the protease induced two-component regulatory system in *B. longum*, we confirmed the regulation of serpin to be different in this species. We first established that unlike in *B. breve*, the regulation of *B. longum* serpin is controlled by different carbohydrates and that this mechanism is conserved in *B. longum* subsp. *longum and B. longum* subsp. *suis* (chapter 3). We showed that the presence of glucose (chapter 3) inhibits the production of serpin in *B. longum* while various other sugars (pentoses, hexoses and oligosaccharides) induce serpin production, albeit to a variable extent. This catabolite-like repression mechanism was first demonstrated in small-scale fermentations (chapter 3) but was further confirmed when NCC 2705 was cultivated in a pH-controlled bioreactor, where glucose depletion allowed the production of serpin (chapter 4). In our initial study, we demonstrated that the absence of glucose was a prerequisite to allow serpin production, but that the presence of galactose or fructose (and related oligosaccharides) was necessary to reach elevated serpin protein levels (chapter 3). In *B. longum* NCC 2705, the glucose/mannose transporter protein (*BL1631*; GlcP) is involved in galactose import [19]. This protein was demonstrated to have the highest specificity for glucose followed by mannose and galactose [17], which could explain why glucose depletion is required to enable the subsequent import of galactose, which once internalized may play a role in the intracellular activation of serpin production. In the same strain, an ATP-binding cassette transporter encoded by the operon *BL0033-0036* (*fruEKFG*) mediates fructose import. Previous studies have shown that the synthesis of the substrate binding protein (FruE) is strongly suppressed by the presence of glucose [43], potentially explaining why in presence of glucose fructose cannot exert its serpin inducing effect.

The above results shed a partial light on the mechanisms controlling serpin production in *B. longum* NCC 2705. In an attempt to better understand the intracellular regulatory circuit involved, we identified in the promoter region upstream of the serpin (*BL0108)* containing operon (i.e., upstream of the *BL0109* gene) a modest resemblance to the AraQ *cis*-acting element and therefore proposed this previously demonstrated global

transcriptional regulator [181] to be implicated in the serpin regulation in *B. longum* NCC 2705 (chapter 3). However, in a subsequent study using a fluorescent-reporter probe we observed that the AraQ controlled promoter P_{BL1359} was only modestly activated when NCC 2705 was grown on galactose and was strongly activated during growth on arabinose (chapter 5). These data are in sharp contrast to our initial results where we demonstrated that serpin levels were only mildly increased on arabinose, while strongly elevated on galactose (chapter 3). Furthermore, we showed that the previously described promoter [137] found upstream of the *BL0109* gene was not differentially activated when the strain was grown on glucose or galactose as sole carbon source, corroborating the fact that similar serpin mRNA levels were detected in both conditions (chapter 7). Overall, these results contradicted the fact that serpin is controlled by AraQ and suggested that serpin might be controlled by a post-transcriptional regulation mechanism in the NCC 2705 strain, which we hypothesized to involve an inhibitory sRNA.

Using transcriptomics analysis of glucose or galactose grown cells, we identified a sRNA encoded between the 16S and 23S rRNA sequences of NCC 2705 that we named sBLO-161. This sRNA displayed an increased expression in glucose grown cells and showed a particularly interesting interaction profile with *BL0108*. It is not only predicted to stably bind to the *BL0108* mRNA by its 5' and 3' ends, but also to form two internal stem-loops that probably inhibit translation of the *BL0108* mRNA (chapter 7). Furthermore, in this study, we highlight the conservation of the 16S-23S rRNA intergenic sequence (and hence sBLO-161) in *B. longum* subsp. *longum* and *B. longum* subsp. *suis* (chapter 7). This is in line with our previous results demonstrating a conservation of the serpin carbohydrate regulation mechanism in these subspecies (chapter 3).

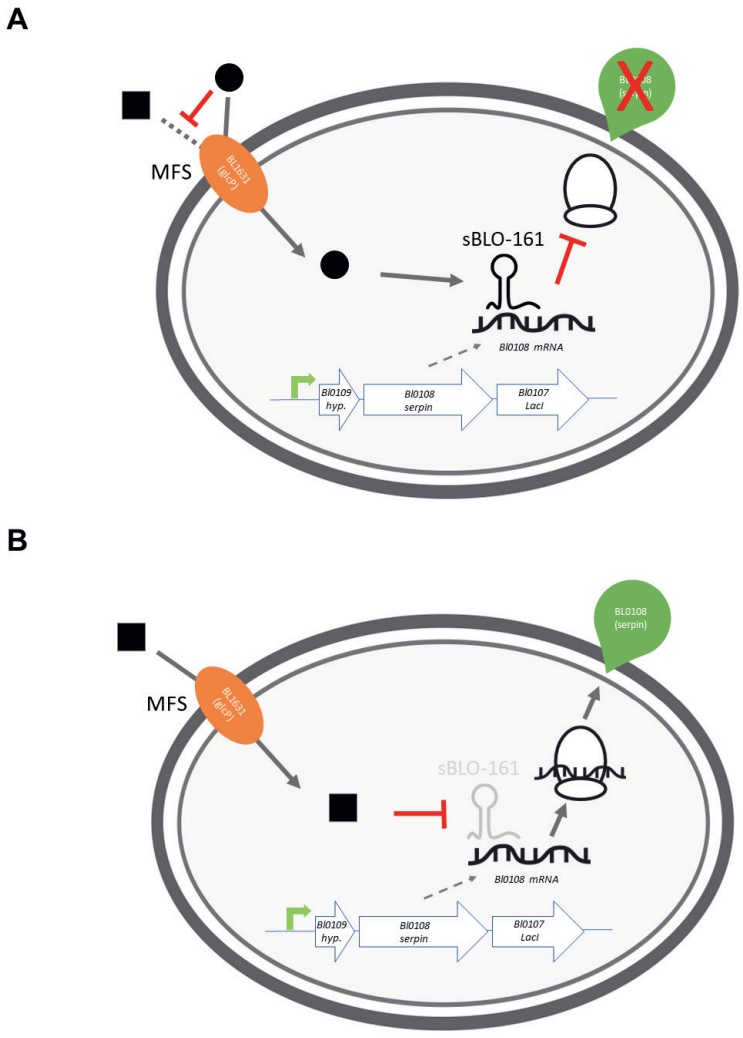


Figure 1 : Serpin (*BL0108*) regulation model in *B. longum* NCC 2705 grown on glucose (panel A) or galactose (panel B). In presence of glucose (black circles), galactose (black squares) import is inhibited, as the glcP (BL1631) MFS protein has a binding preference for glucose. Growth on glucose or galactose does not modify the abundance of the *BL0108* mRNA, but growth on glucose increased expression of the trans-acting sBLO-161 sRNA, which inhibits translation of the *BL0108* mRNA. Absence of glucose enables galactose import, which then decreases sBLO-161 levels, hence allowing BL0108 translation.

Effects and regulation of glucose-catabolite repression in *B. longum* NCC 2705.

Carbohydrate-mediated catabolite repression has been observed in bifidobacteria [177,178] and our studies further support the existence of such a mechanism in *B. longum* NCC 2705. We established that growth on glucose inhibits the production of the serine protease inhibitor (serpin) (chapter 3), which was confirmed in pH-controlled fermentations where glucose depletion allowed the production of serpin (chapter 4).

By performing a detailed comparison of glucose and galactose grown *B. longum* NCC 2705, we found numerous physiological differences between cells grown on either one of these sugars as sole carbon source (chapter 6). Glucose grown NCC 2705 displayed a high growth rate, which was accompanied by several transcriptome metabolic adaptations reflecting this phenotype, i.e., upregulation of nucleotide biosynthesis and ATP generation. Conversely, a reduced growth rate was observed when the strain was grown on galactose, which was accompanied by a smaller cell size and an expanded catabolic flexibility phenotype. The latter was reflected by a variety of transcriptional changes related to carbon source import and central energy metabolism. Galactose growth also modified the metabolic end-products, leading to an increased acetate:lactate ratio and the production of ethanol and formate. We can however not exclude that these changes in the central carbon metabolism fluxes distribution are linked to the metabolization of galactose rather than the absence of glucose. We further observed that galactose grown *B. longum* NCC 2705 displayed an improved cell envelope integrity and enhanced acid-stress resistance as compared to glucose grown cells (chapter 6). By conducting a fully pH-controlled fermentation, the NCC 2705 strain consumed the entire pool of glucose present in the medium during the metabolic active growth phase. In this condition, we noted that upon glucose depletion NCC 2705 cell integrity and acid stress resistance was improved, suggesting that these phenotypes may result from the relief of glucose-catabolite repression (chapter 4). All the above-mentioned results confirm the existence of glucose-controlled catabolite repression similarly to what has been observed in other gram-positive bacteria. However, in some species, catabolite repression is not only mediated by glucose, but also by lactose [176]. It was previously demonstrated that *B. longum* NCC 2705 prefers lactose over glucose [17]. It is therefore plausible that a lactose-mediated catabolite repression mechanism exists in this strain, even if demonstration of this hypothesis requires further studies.

Using a fluorescent-reporter probe system, we next studied how glucose-catabolite repression was regulated in *B. longum* NCC 2705 (chapter 5). We observed a typical diauxic shift when the strain was grown on a mix of glucose and galactose where glucose is consumed first inhibiting the import of galactose, which is only metabolized upon glucose depletion. In contrast, the presence of glucose only partially repressed arabinose utilization, allowing arabinose transport systems to be transcribed at low level and hence enabling both sugars to be co-consumed (chapter 5). These results support the notion that the strength of the glucose-catabolite repression exerted in *B. longum* NCC 2705 is not the same for all alternative sugars. This behavior is particularly relevant for the colonic live style of the bacteria and therefore warrants further studies that aim at deepening our understanding of the underlying regulation mechanisms.

In most Gram-positive bacteria, catabolite repression is controlled by a protein complex of histidine phosphocarrier protein (HPr) and catabolite control protein A (CcpA), a LacI-type regulator that can bind to specific DNA motifs (so-called *cre*-elements) and inhibits or activates transcription of downstream genes [179,222]. Recognizable orthologues of the CcpA protein are absent in bifidobacteria, and typical genome-wide catabolite repression mechanisms have not been described in detail for the members of this genus. Nevertheless, two LacI-type transcriptional regulators, AraQ and MalR1, were shown to control a large set of genes spread over the genome of *B. breve* UCC 2003, regulating a range of transcriptional factors in this strain, and controlling the transcription of genes related to sugar import and metabolism, including genes involved in the ”bifid-shunt” [181,182]. Our results support the involvement of AraQ and MalR1 in the control of glucose-catabolite repression (chapter 6), as out of the 13 genomic loci predicted to be under AraQ transcription control [180], 9 were differentially expressed between galactose and glucose grown cells (Figure 4). In addition, 4 of the 7 loci that were predicted to be regulated by MalR1 in *B. longum* NCC 2705 [180] were confirmed to be induced during growth on galactose, whereas 2 of the MalR1 target loci were induced during growth on glucose (Figure 4) (chapter 6). These findings support the involvement of AraQ and MalR1 in glucose-mediated catabolite repression control. However, the functions regulated by both transcriptional regulators cover only about half of the pleiotropic adaptations observed in galactose grown cells (11

carbohydrate related functions out of the 23 functions we found to be enriched), suggesting that another complementary regulation mechanism may exist.

We observed that elevated serpin protein levels in *B. longum* NCC 2705 were not explained by a significant difference in the corresponding mRNA levels, suggesting the possible implication of a post-transcriptional mechanism (chapter 7). When deciphering the role of sRNAs in the regulation of serpin production, we identified a *trans*-acting sRNA encoded in the intergenic 16S-23S rRNA regions of *B. longum* NCC 2705, which we named sBLO-161. Strikingly, when predicting the potential targets of this sRNA, more than 300 predicted transcript targets were identified in the genome of *B. longum* NCC 2705 (Figure 2). We showed that about 60% of the sBLO-161 targets related to carbohydrate import and metabolism displayed a trustable RNA duplex structure resembling that of the sBLO-161/serpin. This suggests that the initial predicted number of sBLO-161 targets (304) may be overestimated and that a verification of the RNA duplex structures formed should be pursued. Nevertheless, we propose sBLO-161 to play a prominent role in post-transcriptional control of a multitude of transcripts by modifying their translation efficiencies (chapter 7). The sBLO-161 predicted targets encompass a variety of different functional categories that are mostly repressed in cells growing on glucose, implying that the regulatory role of this sRNA may be broader than the combined regulons of AraQ and MalR1. Intriguingly, some genes are predicted to be controlled by both MalR or AraQ transcriptional regulation and post-transcriptional control by sBLO-161 (e.g., several carbohydrate metabolic functions) (Figure 2). Overall, sBLO-161 is predicted to control a large range of functions, including serpin biosynthesis, stress response, cell wall synthesis, sugar transport systems and central carbon metabolism. Within these last two functional categories, sBLO-161 sRNA is foreseen to control not less than 12 different sugar transport systems. Remarkably, for 7 of these targets the predicted RNA duplex structure resembles the sBLO-161/serpin one, including import functions for a variety of sugars (fructose, ribose, xylose, raffinose, galactans, xylans, arabinans, xylosides and galactose) (chapter 7). Moreover, the *trans*-acting sBLO-161 sRNA may also target 9 transcripts of central carbon metabolism functions in NCC 2705 (galactose metabolism, glyceraldehyde-3-phosphate dehydrogenase, 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase, and xylose/fructose-6-

phosphoketolase enzyme), including the F6PPK transcript that plays a pivotal role in the bifid-shunt (chapter 7).

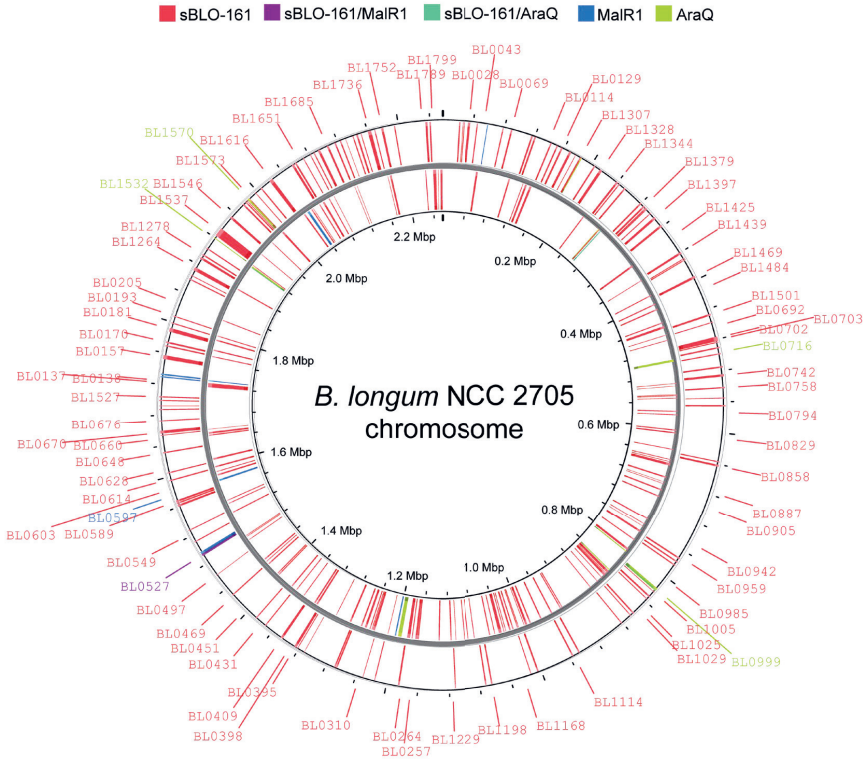


Figure 2 : Predicted sBLO-161 (red), AraQ (green) and MalR1 (blue) regulated genes mapped on the circularized chromosome of *B. longum* NCC 2705 [16]. The sBLO-161 regulons were identified using RNApredator (chapter 7) while the AraQ and MalR1 predicted regulons were obtained from the Regprecise database. The map was constructed using the online available Proksee tool (*https://proksee.ca/*) derived from the cgview tool [283].

A gene ontology enrichment performed using all direct (chapters 6 and 7) and indirect (Table 1) AraQ, MalR1 and sBLO-161 regulons showed that the functions controlled by these transcriptional and post-transcriptional regulatory circuits (Figure 3) recapitulated relatively well the gene ontology enrichment performed on the differentially expressed genes obtained from the galactose grown culture (chapter 6). This confirms the implication of these three regulatory elements in most of the phenotypic changes observed in galactose grown cells, compared to their glucose counterparts (chapter 6). More specifically, it argues for a role of AraQ, MalR1 and sBLO-161 in the control of cell envelop synthesis (amino-sugar and cellular polysaccharides synthesis) and carbohydrate metabolism (e.g., monosaccharides, hexoses, galactose, glucose and alcohol metabolic processes).

Table 1 : List of the different regulatory proteins regulated by the Lac-I type regulators AraQ and malR1 or the trans-acting sRNA sBLO-161. When known, the predicted regulated functions and set of predicted regulons are indicated.

Gene Id	Gene name	Uniprot annotation	Regulated by	Predicted regulated function	Predicted regulon
BL0275	AraQ	LacI-type transcriptional regulator	AraQ	autoregulation	BL0275
BL0528	MalR1	probable LacI-type transcriptional regulator	MalR1	autoregulation	BL0528
BL0107		Probable LacI-type transcriptional regulator	sBLO-161	Sucrose import & metabolism	BL0105-BL0106
BL0185		LacI-type transcriptional regulator	sBLO-161	Arabinoxylans degradation, import and metabolism	BL0183-BL0187
BL0289		MerR-type transcriptional regulator	sBLO-161	Unkown in NCC 2705	
BL0426		LacI-type transcriptional regulator	sBLO-161	Arabinose transport & metabolism	BL0422-BL0426
BL0516		Probable MerR-type-transcriptional regulator	sBLO-161	Unkown in NCC 2705	
BL0718	hrcA	Heat-inducible transcription repressor HrcA	sBLO-161	Heat shock response	BL002, BL0718-BL0719, BL1558
BL0786	cpsY	LysR-type transcriptional regulator	sBLO-161	Unkown in NCC 2705	
BL1000		Response regulator of two-component system	sBLO-161, AraQ	unkown, part of a two component system coupled to a GH13 type branching enzyme	
BL1128	fur	Probable metal uptake regulator similar to ferric uptake regulator protein	sBLO-161	Zinc homeostasis	BL1126; BL1309
BL1199A	moxR	Regulatory protein homolog	sBLO-161	Methanol dehydrogenase regulatory protein homolog	
BL1212		Probable DeoR-type transcriptional regulator	sBLO-161	Galactose metabolism	BL1210-BL1212; BL1671
BL1493	whiB	WhiB-like family protein	sBLO-161	Multiple stress response (heat, osmotic, oxidative, starvation)	unkown in NCC 2705

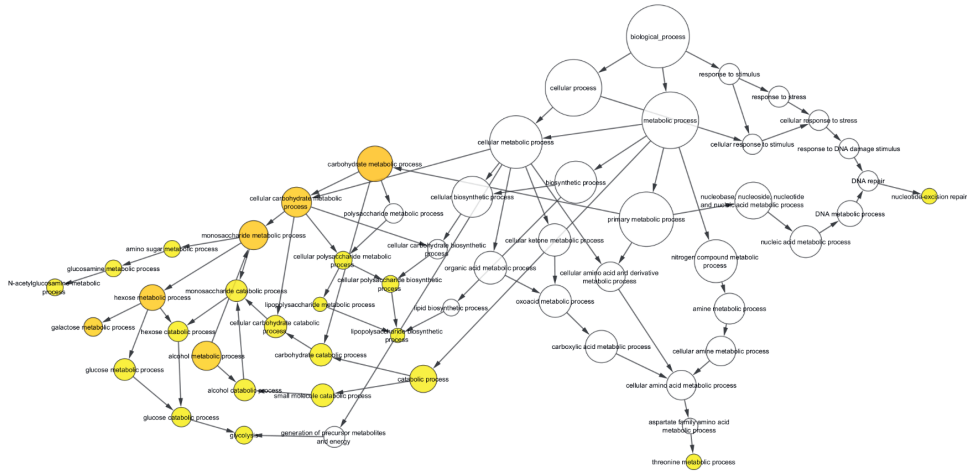


Figure 3 : Gene Ontology (GO) functions enriched in the AraQ, MalR1 and sBLO-161 regulated genes. This analysis used a set of directly (333) and indirectly (21) controlled genes. GO enrichment analysis was performed using the Cytoscape BINGO plugin, using the hypergeometric statistical test. Enriched functions (p-value <0.05) are displayed on a yellow-based scale.

To better understand the interplay between the three regulatory molecules in orchestration of carbohydrate metabolism of *B. longum* NCC 2705, we attempted to reconstruct a “cellular map” representing the overall carbohydrate import and metabolism repertoire (Figure 4). This analysis revealed the complementary role of AraQ, MalR1 and the *trans*-acting sRNA sBLO-161. Overall, not less than 38 genes are predicted to be directly regulated by these three regulatory elements, which represents almost half (43%) of the genes possibly involved in the carbohydrate import and metabolism of *B. longum* NCC 2705 (Figure 4). They are furthermore predicted to regulate the expression of a set of 14 regulatory proteins (Table 1) including their own autoregulation, thus extending their indirect regulation role to other carbohydrate related functions like sucrose import and metabolism (through the LacI-type regulator BL0107), arabino-xylan/galactan degradation and metabolism (through the LacI-type regulators BL0185 and BL0426, as well as the DeoR-type regulator BL1212), or functions related to stress responses (through BL0718 [HrcA] and BL1493 [WhiB]). Altogether, we propose AraQ, MalR1 and sBLO-161 to act in concert and participate in the control of glucose-catabolite repression in *B. longum* NCC 2705.

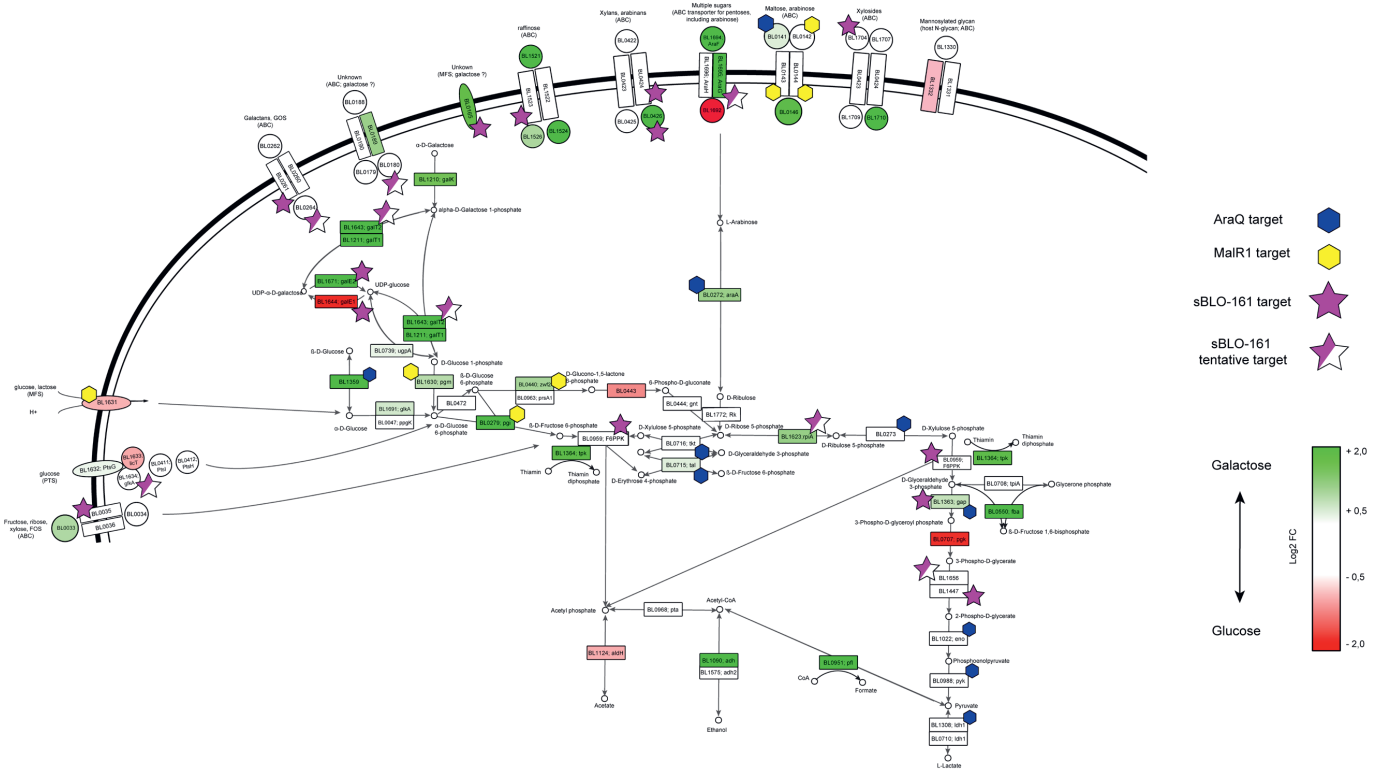


Figure 4 : Model recapitulating the proposed catabolite-regulation mechanism controlling the carbohydrate import and metabolism of *B. longum* subsp. *longum* NCC 2705. The different sugar import systems and central carbon metabolism ("bifid-shunt") implicated genes are represented. The color scale represents the differential expression (log2) measured by transcriptomics in glucose (colored in red) or galactose (colored in green) grown cells. Genes that are predicted to be transcriptionally controlled by AraQ or MalR1 (as predicted by Regprecise) are indicated using a blue or yellow polygon, respectively. Proteins for which the translation is predicted to be controlled by the trans-acting sBLO-161 are highlighted by a purple star. Full purple stars represent mRNA for which sBLO-161 was shown to form a stable RNA-RNA dimer, while the ones with a tentative binding profile are depicted using a semi-colored purple star.

Deciphering the regulation of sBLO-161

We showed that sBLO-161 had a consistent size (168-9 bp) in all evaluated samples and that its expression was increased in the fast glucose growing *B. longum* NCC 2705 as compared to its galactose slow growing counterpart (chapter 7). It remains, however, unknown how sBLO-161 is exactly produced and regulated. Bacterial growth laws were formulated based on knowledge gathered in the model organism *E. coli* [284]. In particular, a remarkable linear relation has been observed between growth rate and total RNA and protein ratios, which are directly proportional to the mass fraction of ribosomes in the cell [284]. These bacterial growth laws are meant to be applicable to all prokaryotes [285] and imply the presence of a higher ribosomal RNA fraction in fast growing cells. This theory hence suggests an increased presence of the 16S-23S polycistronic RNA in fast growing cells (e.g., glucose grown *B. longum* NCC 2705), which we propose to be the precursor of sBLO-161. Further demonstration of the correlations between growth rate, ribosomal RNA fraction and finally sBLO-161 and its regulated functions (e.g., serpin, catabolic flexibility) would support the role of this sRNA as a global growth rate regulator controlling a wide range of associated functions in bifidobacteria.

We hypothesize sBLO-161 levels to be governed by the ribosomal RNA fraction in *B. longum* NCC 2705 and propose this *trans*-acting sRNA to be produced by splicing of the polycistronic 16S-23S rRNA. In the model organism *E. coli*, the polycistronic 16S-23S rRNA was shown to be processed by a set of endo- and exo- ribonucleases (RNases) to generate mature and functional rRNAs. Maturation is initiated by RNase III, which cleaves the rRNA transcript to generate precursors of the three 16S, 23S and 5S rRNAs [286]. The 3' end of the 16S precursor rRNA is then matured by the action of multiple exoribonucleases (RNases II, R, PH, E and PNPase) [287] while the 5' end of the 23S precursor rRNA is matured by RNase AM, that actually acts on all three ribosomal RNAs 5' ends [288]. The maturation sequence and the involved ribonucleases are unknown to date in bifidobacteria. Homologues of some of the above-mentioned RNases are known in the genome of *B. longum* NCC 2705 (*BL0295* [*RNase III*], *BL0286* [*rph*], *BL1546* [*PNPase*]), and the BLASTp analysis we performed may have further identified homologues of *RNase E* (*BL1281*, annotated as ribonuclease G [39,8% coverage, 35% ID]) and *RNase AM* (*BL0089*, annotated as POLIIIAc domain-containing protein [95,5% coverage, 30% ID]). Furthermore, several endonuclease

encoding genes were predicted to be translationally controlled by sBLO-161 (*BL0703; BL0702; BL1025; BL0393; BL1812; BL1207),* which may suggest that its production is autoregulated.

Digging into our transcriptomics data (chapter 6), most of the above endonucleases showed relatively low transcriptional levels (Figure 5). For some of those genes, the numbers were even below the threshold used in our analysis (i.e., <5 counts in at least two samples) and were hence initially excluded as "not-expressed" (*BL0703, BL1207*). Re-analysis of the normalized transcript levels showed that the *BL0089* encoded RNase was not only expressed at higher levels but was upregulated when the strain was grown on glucose as compared to galactose. We predict this gene to be a homologue of RNase AM, which was previously demonstrated in *E. coli* to be responsible of the 5'-end maturation of the different ribosomal sRNA [288], including the 5' end of the 23S rRNA which may correspond to the 3' end of sBLO-161. Overall, the implication of the above ribonucleases (and particularly the *BL0089* encoded one) in the production sBLO-161 should be examined further.

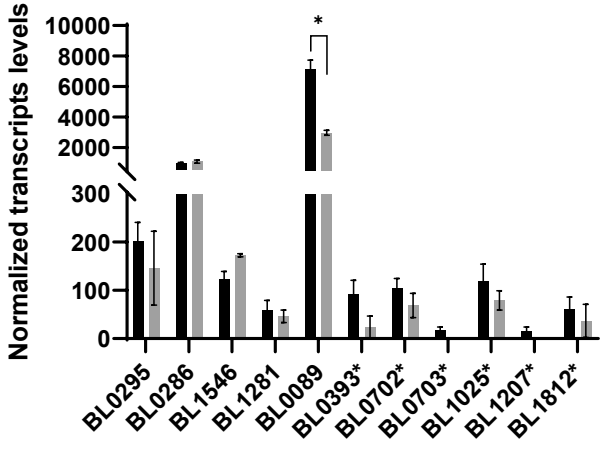


Figure 5 : Transcriptional levels of *B. longum* NCC 2705 ribonucleases which may play role in the production of sBLO-161, including homologues of the ribonucleases previously shown to be involved in the maturation of the 16S or 23S rRNA transcripts in *E. coli* and ribonucleases predicted to be regulated by sBLO-161 (indicated with a *). Averaged transcript levels found in glucose (black bars) or galactose (grey bars) grown *B. longum* NCC 2705 are displayed with their standard deviation. Statistical analysis was performed using an unpaired t test with Welch multiplicity correction; * q-value <0.05.

The intrinsic challenges related to sRNA expression engineering

In an attempt to clarify the role of sBLO-161 in the regulation of the serpin production in *B. longum* NCC 2705, we tried and failed to overexpress this *trans*-acting sRNA by heterologous expression. Our observations suggested that the proposed global regulatory role of sBLO-161 prevents its overexpression due to the pleiotropic effects of cellular physiology this could cause (chapter 7).

However, we cannot exclude that these results arise from our cloning strategy. First, we selected a constitutive promoter (P_{gap}) which is very strong in NCC 2705 [289]. As we propose sBLO-161 to control a large set of functions in this strain, it is plausible that producing a large amount of this regulatory sRNA is causing detrimental effects to the cells. The choice of weaker (e.g., P_{BL1518}) or inducible (e.g., P_{BL1359} or P_{BL1694}) promoters could then be a more suitable alternative (chapter 5). Small RNAs are small single stranded RNAs, and as opposed to mRNA molecules, they do not harbor the canonical translation signals (start and stop codons) at their 5'- and 3'- ends that define the encoded protein when the mRNA is translated by ribosomes. This denotes a fundamental difference between sRNA and mRNA and represents a real challenge in engineering sRNA overexpressing plasmids, as unintended extensions of the sRNA molecule at its 5'- or 3'- end can likely interfere with its target recognition and affinity. In the case of sBLO-161 this is quite obvious based on its 5'- and 3'- end involvement in mRNA target recognition. Engineering the 3'-end of a sRNA transcript is particularly delicate because transcription termination has to take place with high accuracy and the prediction of the precise position of transcription termination using a terminator sequence is far from trivial. To summarize, it is more complicated to engineer overexpression of a sRNA as compared to a mRNA, which may also have hampered our attempts to overproduce sBLO-161. Overall, this suggests that a heterologous overexpression strategy for sRNA molecules is not optimal and deciphering the regulatory role of these small RNA molecules may be better approached using alternative strategies. One of them could consist in modifying the (specific 5'- and 3'-end) sequence of sBLO-161. This could be achieved using a "gene"-replacement strategy, although considering there are multiple copies of the rRNA locus, all of these would need to be modified to ensure a precise intervention in the target-binding activity of the resulting sRNA. Moreover, sequence changes in the intergenic region of the 16S and 23S genes may affect the production of the adjacent 16S and 23S rRNA

sequences and the ribosomes biogenesis, which will have detrimental consequences. An attractive approach could be to employ one of the recent gene-editor tool based on CRISPR-Cas systems that would allow specific single nucleotide modification. This approach may leverage bifidobacterial-adapted systems that several research groups are trying to develop [290,291].

Finally, the sBLO-161 sRNA appears at least partially conserved amongst *B. longum* subspecies (chapter 7) and may even be conserved amongst other species of *Bifidobacterium*, warranting further investigation of the regulatory role of this sRNA in the members of these species and genus. It is intriguing that sBLO-161 is encoded in one of the most conserved genomic loci in the bacteria world and specifically in an intergenic region that has been used for species or subspecies delineation in different studies [280,281]. Our analysis detected a substantial level of conservation of the sBLO-161 sequence amongst several *B. longum* subspecies, but also established sequence divergence in *B. longum* subsp*. infantis* particularly in the 5'-end that is predicted to play a key-role in the formation of the sRNA-mRNA duplex (chapter 7). Therefore, sBLO-161 homologues could also participate in post-transcriptional regulation in other *B. longum* subspecies and could even be conserved amongst other species of *Bifidobacterium*. Future investigations aiming at clarifying the existence of sBLO-161 homologues in other strains and species could be of interest to decipher the potentially generic role of these sRNAs in members of the *Bifidobacterium* genus or even beyond, in bacteria in general.

Developments of *B. longum* serpin-based probiotic concepts

Growing evidence associates imbalanced proteolytic activity in the digestive tract to the development of inflammatory diseases [292]. Several studies have demonstrated an elevated proteolytic activity (e.g., excessive serine protease activity) in intestinal tissues and fecal samples of patients suffering from Inflammatory Bowel Diseases (IBD) [293,294] or food triggered immune disease such as celiac disease [295]. Excessive proteolysis leads to a disruption of the gut barrier, via cleavage of intercellular junction proteins or opening of the tight junction through protease activated receptors. In addition, proteases regulate the activity and availability of cytokines and growth factors, which are also known to modulate intestinal permeability and participate in exacerbated gut inflammation [296,297].

Following these observations, therapeutic strategies that aim at regulating gut serine proteases activity were proposed to treat and control the above-mentioned inflammatory diseases [297,298]. Serine protease inhibitors (serpins) comprise the most widely distributed superfamily of protease inhibitors and possess demonstrated anti-inflammatory properties [299]. It was for example demonstrated that transgenic mice expressing the human serpin elafin, a tissue-derived inhibitor of elastase, were protected from colitis. The same study demonstrated that a recombinant *L. lactis* producing elafin could decrease elastolic activity and inflammation, thus restoring intestinal homeostasis in a mouse model of acute colitis [300]. The same recombinant strain was demonstrated to elicit a similar beneficial effect in a mouse model of celiac disease [295], demonstrating that an elafin-delivering bacteria can effectively protect from the onset of those inflammatory diseases.

The initial study revealing the presence of a serpin encoding gene in NCC 2705 demonstrated that a recombinantly produced version of this protein could inhibit both pancreatic and neutrophil elastases [61]. Inhibition of neutrophil elastase, which is a driver of intestinal tissue damage and a biomarker of intestinal inflammation [301], represents a particularly relevant mean of managing gastrointestinal inflammatory conditions. This observation supports the hypothesis that the NCC 2705 serpin protein is a functional homologue of elafin and hence justifies the development of *B. longum* NCC 2705 serpin-based concepts for the prevention or treatment of (a range of) inflammatory diseases. Below, we propose two possible developments that build on

the findings in this thesis, where we discuss possible therapeutic solutions that target duodenal or colonic inflammatory diseases and could open the way to future human clinical trials demonstrating a consistent protective effect of *B. longum* NCC 2705.

***B. longum* NCC 2705 with enhanced serpin levels as adjuvant therapy for celiac disease patients.** Celiac Disease (CD) is a chronic systemic immune-mediated disorder triggered by gluten consumption, which primarily affects the small intestine of individuals who are genetically predisposed [302-304]. The best-characterized genetic risk factor for CD is the presence of genes encoding for Major Histocompatibility Complex (MHC) class II antigen proteins (HLA-DQ 2 and HLA-DQ8) that are found at the surface of human leukocytes [303]. The dietary trigger, gluten, comprises various proteins found in wheat and other related grains such as barley, and rye [305]. In the small intestine gluten immunogenic peptides produced during digestion by proteolysis (in particular a specific 33-mer) cross the loosened epithelial barrier, are de- and trans-amidated by tissue transglutaminase (TTG) and subsequently presented by antigen presenting cells (HLA-DQ2/DQ8 positive cells), inducing a CD4+ T-lymphocyte response, which in turn leads to high-levels of pro-inflammatory cytokines [303,304,306].

It has previously been demonstrated in a mouse model of celiac disease that NCC 2705 could dampen gluten-induced immunopathology [64]. This protective effect was shown to be equivalent to that of the elafin-producing recombinant *L. lactis* [295] and was demonstrated to be serpin-mediated, since a serpin knock-out derivative of NCC 2705 failed to protect the sensitized mice [64]. Recently, a safety study performed in celiac disease and non-celiac gluten sensitive patients revealed that *B. longum* NCC 2705, delivered at a daily dose of 1E10 CFU is well tolerated and safe for human consumption. In the same study, NCC 2705 was detected in duodenal aspirates at a level of 6.7 log copies genomes/ml peaking at 90 minutes after ingestion. Strikingly, the maximal concentration of NCC 2705 serpin, measured using the serpin specific ELISA (chapter 3), was reached at the same time point (concentration between 4.9 pg/ml to 24.5 pg/ml) [68]. Overall, the above-mentioned results provide a mechanistic basis for the potential beneficial use of the strain in a celiac disease population.

We and others have suggested a possible role of manufacturing conditions in some of the discrepancies observed in clinical trials using probiotics [101,102,116] (chapter 2). We advocate that it is important to ensure the presence of molecules driving the health beneficial effect of the strain in the final probiotic product. We consider this particularly relevant to treat or prevent proximal intestinal tract located diseases, since the relatively short transit time to reach the duodenum following consumption is unlikely to allow molecular adaptation of the probiotic bacteria (chapter 2). Celiac disease is

primed and elicited by gluten exposure in the proximal small intestine, hence it appears important to ensure that maximal amounts of serpin are found at the surface of *B. longum* NCC 2705 when using the strain to dampen symptoms in celiac disease patients.

Findings of this thesis can be leveraged to develop improved probiotic production processes aiming at maximizing the levels of serpin exposed at the surface of *B. longum* NCC 2705. For example, an alternative carbon source (chapter 3) or an adapted pH-regulation process (chapter 4) could be used to produce a serpin-rich *B. longum* NCC 2705 biomass. Moreover, the ELISA we developed (chapter 3) can be used to ensure that the downstream processing does not impair the presence of the molecule at the surface of the bacteria. The effect of the biomasses obtained by the classical and modified production processes could further be compared in the mouse model of celiac disease [64] to support the above proposal. One would expect an improved effect of the serpin enriched NCC 2705 preparation or a reduced required dosage to achieve the effect. Ensuring the presence of the serpin molecule at the surface of the bacteria upon manufacturing may become an important factor to check in the final probiotic preparations beyond the traditional live cells (e.g., CFU) enumeration. Such effector molecule-oriented production process could significantly contribute to establish the consistent efficacy of NCC 2705 preparations in attenuating the immunological effects induced by gluten in celiac disease patients.

A tailored *B. longum* NCC 2705 synbiotic for the prevention of IBD. As evoked above, data support the *B. longum* serpin to be a functional homologue of the human elafin and as such to play a role in intestinal homeostasis, similarly to the recombinant *L. lactis* strain producing elafin [61,64,300].

This is what was demonstrated by Foligné and colleagues in a mouse model of acute colitis (unpublished data, personal communication). The strains previously used in the murine celiac-model study [64], i.e., *B. longum* NCC 2705 and a serpin knock-out (Δserpin) derivative, as well as a constitutively expressing (serpin$^+$) derivative were administered at a dose of 1E8 CFU/day before induction of the experimental colitis. Three days after onset of the colitis, development of the disease was assessed by macroscopic (Wallace score) and histological evaluations. As shown in Figure 6 (courtesy of Foligné et al), the Wallace score (Figure 6 A) commonly used to assess

colonic inflammation (the stronger the inflammation, the highest the score) was clearly lower in mice treated with the wild type *B. longum* NCC 2705 strain as compared to untreated controls. Notably, the protective effect was lost in mice treated with the Δserpin strain but enhanced after treatment with a strain constitutively producing serpin (serpin^{+}). These differential effects also translated remarkably in histological examination of colonic tissues (Figure 6 B). These data clearly support the key role played by serpin in the anti-inflammatory effect of NCC 2705 and underline the importance of the serpin levels delivered in the mice gut.

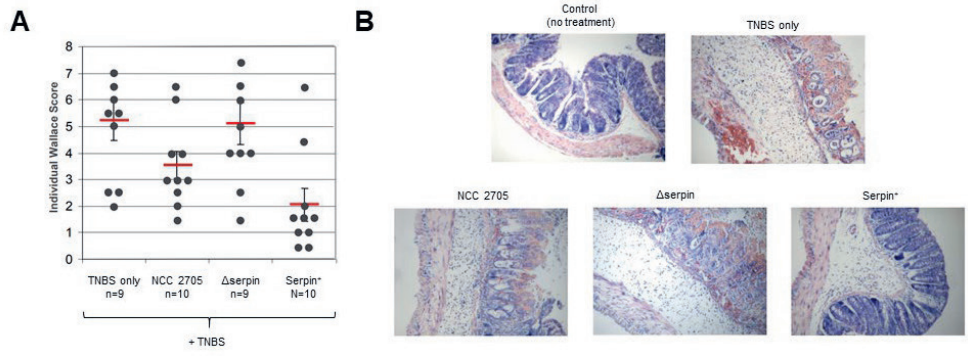


Figure 6 : Courtesy of Foligné et al : effect of wild type *B. longum* NCC 2705, a serpin knock-out (Δserpin) and a constitutively expressing (serpin+) recombinant strain in a mouse model of acute colitis. Macroscopic evaluations (Wallace score) (panel A) and histology of colon tissues (panel B) are represented.

To maximize the effect of serpin-rich NCC 2705 biomass, the targeted site of action of the probiotic strain has to be taken into account. IBD affects primarily the human colon, as opposed to celiac disease which predominantly affects the duodenum. Therefore, NCC 2705 should ideally deliver serpin in the colon, facing a longer transit time upon consumption, which may allow the strain to undergo molecular adaptations impacting the cell surface associated serpin levels. Therefore, production of a serpin enriched probiotic preparation may be less effective for the management of IBD (as compared to celiac disease). Nevertheless, this thesis provides avenues to enhance serpin expression of *B. longum* NCC 2705 in the colon. To this end, we propose to co-administer the strain with a tailored prebiotic (i.e., a sugar) that is known to induce serpin expression.

The current prebiotic definition is “a substrate that is selectively utilized by host microorganisms conferring a health benefit” [307]. According to this definition, the

selected prebiotic will have to be undigestible by humans and selectively used by NCC 2705 to ensure sufficient serpin delivery *in situ*. Preferably, the selected carbohydrate should contain a significant amount of galactose or fructose, which we have shown to increase serpin production in NCC 2705 (chapter 3). Several prebiotic ingredients may serve that purpose, but the human milk oligosaccharide Lacto-N-teatraose (LNT; 50 % galactose moieties), inulin or its derived fructo-oligosaccharides (FOS, e.g., Orafti P95; containing a high percentage of branched-fructose units) may represent suitable candidates. It is known that LNT is undigestible by humans [169] and is currently considered as a relatively specific substrate for bifidobacteria, supporting the growth of *B. longum* [140] and *B. breve* [308]. However, it remains largely unknown to what extent LNT can be consumed by other gut commensals, and a better understanding of the selectivity of this HMO in stimulation of specific bifidobacteria is needed.

Also, inulin or the derived FOS were demonstrated to increase bifidobacterial abundance in the gut. FOS can be metabolized by NCC 2705 [19], while the higher molecular weight inulin substrate is not digested by this strain [78]. Nevertheless, despite the lack of metabolization by NCC 2705 inulin was shown to specifically increase the abundance of bifidobacteria, with a prominent increase of *B. adolescentis* followed by *B. longum*, where the latter species likely benefits from the initial degradation of inulin by *B. adolescentis* [309]. FOS may likely serve as a substrate for several members of the gut microbiota, possibly leading to competition or cooperation with other intestinal species, which remains to be better studied. Further studies, first conducted *in vitro* (e.g., fecal fermentation experiments), should investigate which is the best and most specific *in situ* serpin enhancing prebiotic. This should then further be followed by *in vivo* experiments confirming that, at least in rodents, the synbiotic concept [310] can be demonstrated. The use of the wild type and serpin knock-out strains in the selected murine model would underpin the role of serpin in the observed effects.

We propose that such a rational synbiotic design might improve the consistency of the serpin mediated health effects delivered by *B. longum* NCC 2705. This concept may as well benefit other target populations suffering from chronic inflammatory conditions characterized by excessive serine proteolytic activity in the colon.

Concluding remarks

The work performed in this thesis significantly advanced our understanding of serpin regulation in *B. longum* subsp. *longum*. Not only did we discover environmental triggers that control the production of this important effector molecule in *B. longum* (chapter 3), but we also revealed that this protein is not controlled through a usual transcriptional control mechanism (chapter 7). We demonstrated how these learnings can be used to drive an optimization of *B. longum* NCC 2705 fermentation processes (chapter 4), moving towards production scenarios that integrate functional probiotic aspects on top of the standard yield outcomes that are commonly used in industry today (chapter 2).

Beyond the initial objectives of this thesis, we contributed to the overall understanding of bifidobacterial physiology. First, we demonstrated the existence of glucose-catabolite repression in bifidobacteria (chapters 5 and 6), for which only partial demonstration was available before our work. Our data highlight the relatively wide range of consequences this mechanism has on the physiology of the *B. longum* NCC 2705 (chapter 6), some of them being particularly relevant for industrial applications (e.g., enhanced stress resistance, increased serpin production). Furthermore, we shed light on the regulation governing the glucose-mediated catabolite repression, confirming the involvement of previously identified global transcriptional regulators (AraQ and MalR1) (chapter 6) and identifying for the first time in bifidobacteria the implication of a *trans*-acting sRNA with a global regulatory role (sBLO-161) (chapter 7). Particularly, the presence and role of this trans-acting regulatory sRNA adds a significant piece of information to our understanding of bifidobacterial physiology, highlighting the importance that this molecule can have in controlling a range of biological functions in this genus.

In conclusion, we feel that this thesis underlines the need to understand fundamental aspects of probiotic physiology. Beyond the necessity to unravel their mode of action in the targeted health benefit(s), we highlight the importance of understanding how the production of effector molecules is regulated. We believe that these basic requirements need to be fulfilled to develop effect-oriented, robust and reproducible probiotic products for specific health applications.

References

1 Tissier, H. La reaction cromophile d'Escherichia et le bacterium coli. *Com rend Soc Biol*, 943-945 (1899).

2 Tissier, H. Traitement des infections intestinales par la méthode de transformation de la flore bacteriéne de l'intestin. *Comptes rendus des séances de la Société de Biologie et de ses filiales* 50, 359 (1907).

3 Alessandri, G., van Sinderen, D. & Ventura, M. The genus bifidobacterium: From genomics to functionality of an important component of the mammalian gut microbiota running title: Bifidobacterial adaptation to and interaction with the host. *Comput Struct Biotechnol J* 19, 1472-1487, doi:10.1016/j.csbj.2021.03.006 (2021).

4 Mattarelli, P. *et al.* Recommended minimal standards for description of new taxa of the genera Bifidobacterium, Lactobacillus and related genera. *Int J Syst Evol Microbiol* 64, 1434-1451, doi:10.1099/ijs.0.060046-0 (2014).

5 Lugli, G. A. *et al.* Investigation of the evolutionary development of the genus Bifidobacterium by comparative genomics. *Appl Environ Microbiol* 80, 6383-6394, doi:10.1128/AEM.02004-14 (2014).

6 Hidalgo-Cantabrana, C. *et al.* Bifidobacteria and Their Health-Promoting Effects. *Microbiol Spectr* 5, doi:10.1128/microbiolspec.BAD-0010-2016 (2017).

7 Avershina, E. *et al.* Bifidobacterial succession and correlation networks in a large unselected cohort of mothers and their children. *Appl Environ Microbiol* 79, 497-507, doi:10.1128/AEM.02359-12 (2013).

8 Garrido, D., Dallas, D. C. & Mills, D. A. Consumption of human milk glycoconjugates by infant-associated bifidobacteria: mechanisms and implications. *Microbiology* 159, 649-664, doi:10.1099/mic.0.064113-0 (2013).

9 Zivkovic, A. M., German, J. B., Lebrilla, C. B. & Mills, D. A. Human milk glycobiome and its impact on the infant gastrointestinal microbiota. *Proc Natl Acad Sci U S A* 108 Suppl 1, 4653-4658, doi:10.1073/pnas.1000083107 (2011).

10 Duboux, S., Ngom-Bru, C., De Bruyn, F. & Bogicevic, B. Phylogenetic, Functional and Safety Features of 1950s B. infantis Strains. *Microorganisms* 10, 203 (2022).

11 Turroni, F. *et al.* Genome analysis of Bifidobacterium bifidum PRL2010 reveals metabolic pathways for host-derived glycan foraging. *Proc Natl Acad Sci U S A* 107, 19514-19519, doi:10.1073/pnas.1011100107 (2010).

12 Egan, M. *et al.* Cross-feeding by Bifidobacterium breve UCC2003 during co-cultivation with Bifidobacterium bifidum PRL2010 in a mucin-based medium. *BMC Microbiol* 14, 282, doi:10.1186/s12866-014-0282-7 (2014).

13 Egan, M., O'Connell Motherway, M., Ventura, M. & van Sinderen, D. Metabolism of sialic acid by Bifidobacterium breve UCC2003. *Appl Environ Microbiol* 80, 4414-4426, doi:10.1128/AEM.01114-14 (2014).

14 Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* 486, 222-227, doi:10.1038/nature11053 (2012).

15 Turroni, F. *et al.* Exploring the diversity of the bifidobacterial population in the human intestinal tract. *Appl Environ Microbiol* 75, 1534-1545, doi:10.1128/AEM.02216-08 (2009).

16 Schell, M. A. *et al.* The genome sequence of Bifidobacterium longum reflects its adaptation to the human gastrointestinal tract. *Proc Natl Acad Sci U S A* 99, 14422-14427, doi:10.1073/pnas.212527599 (2002).

17 Parche, S. *et al.* Lactose-over-glucose preference in Bifidobacterium longum NCC2705: glcP, encoding a glucose transporter, is subject to lactose repression. *J Bacteriol* 188, 1260-1265, doi:10.1128/JB.188.4.1260-1265.2006 (2006).

18 Kitaoka, M., Tian, J. & Nishimoto, M. Novel putative galactose operon involving lacto-N-biose phosphorylase in Bifidobacterium longum. *Appl Environ Microbiol* 71, 3158-3162, doi:10.1128/AEM.71.6.3158-3162.2005 (2005).

19 Parche, S. *et al.* Sugar transport systems of Bifidobacterium longum NCC2705. *J Mol Microbiol Biotechnol* 12, 9-19, doi:10.1159/000096455 (2007).

20 Duranti, S. *et al.* Genomic characterization and transcriptional studies of the starch-utilizing strain Bifidobacterium adolescentis 22L. *Appl Environ Microbiol* 80, 6080-6090, doi:10.1128/AEM.01993-14 (2014).

21 O'Callaghan, A. & van Sinderen, D. Bifidobacteria and Their Role as Members of the Human Gut Microbiota. *Front Microbiol* 7, 925, doi:10.3389/fmicb.2016.00925 (2016).

22 Zannini, E., Bravo Nunez, A., Sahin, A. W. & Arendt, E. K. Arabinoxylans as Functional Food Ingredients: A Review. *Foods* 11, doi:10.3390/foods11071026 (2022).

23 Hiel, S. *et al.* Effects of a diet based on inulin-rich vegetables on gut health and nutritional behavior in healthy humans. *Am J Clin Nutr* 109, 1683-1695, doi:10.1093/ajcn/nqz001 (2019).

24 Kleessen, B. *et al.* Jerusalem artichoke and chicory inulin in bakery products affect faecal microbiota of healthy volunteers. *Br J Nutr* 98, 540-549, doi:10.1017/S0007114507730751 (2007).

25 Turroni, F. *et al.* Deciphering bifidobacterial-mediated metabolic interactions and their impact on gut microbiota by a multi-omics approach. *ISME J* 10, 1656-1668, doi:10.1038/ismej.2015.236 (2016).

26 Fukuda, S. *et al.* Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature* 469, 543-547, doi:10.1038/nature09646 (2011).

27 De Vuyst, L. & Leroy, F. Cross-feeding between bifidobacteria and butyrate-producing colon bacteria explains bifdobacterial competitiveness, butyrate production, and gas production. *International Journal of Food Microbiology* 149, 73-80, doi:https://doi.org/10.1016/j.ijfoodmicro.2011.03.003 (2011).

28 Blaak, E. E. *et al.* Short chain fatty acids in human gut and metabolic health. *Benef Microbes* 11, 411-455, doi:10.3920/BM2020.0057 (2020).

29 Tan, J. *et al.* The role of short-chain fatty acids in health and disease. *Adv Immunol* 121, 91-119, doi:10.1016/B978-0-12-800100-4.00003-9 (2014).

30 van der Hee, B. & Wells, J. M. Microbial Regulation of Host Physiology by Short-chain Fatty Acids. *Trends Microbiol* 29, 700-712, doi:10.1016/j.tim.2021.02.001 (2021).

31 Sonnenburg, J. L., Chen, C. T. & Gordon, J. I. Genomic and metabolic studies of the impact of probiotics on a model gut symbiont and host. *PLoS Biol* 4, e413, doi:10.1371/journal.pbio.0040413 (2006).

32 Riviere, A., Gagnon, M., Weckx, S., Roy, D. & De Vuyst, L. Mutual Cross-Feeding Interactions between Bifidobacterium longum subsp. longum NCC2705 and Eubacterium rectale ATCC 33656 Explain the Bifidogenic and Butyrogenic Effects of Arabinoxylan Oligosaccharides. *Appl Environ Microbiol* 81, 7767-7781, doi:10.1128/AEM.02089-15 (2015).

33 Wang, Y. & LaPointe, G. Arabinogalactan Utilization by Bifidobacterium longum subsp. longum NCC 2705 and Bacteroides caccae ATCC 43185 in Monoculture and Coculture. *Microorganisms* 8, doi:10.3390/microorganisms8111703 (2020).

34 Salazar, N. *et al.* Exopolysaccharides produced by Bifidobacterium longum IPLA E44 and Bifidobacterium animalis subsp. lactis IPLA R1 modify the composition and metabolic activity of human faecal microbiota in pH-controlled batch cultures. *Int J Food Microbiol* 135, 260-267, doi:10.1016/j.ijfoodmicro.2009.08.017 (2009).

35 Li, S. *et al.* The beneficial effect of exopolysaccharides from Bifidobacterium bifidum WBIN03 on microbial diversity in mouse intestine. *J Sci Food Agric* 94, 256-264, doi:10.1002/jsfa.6244 (2014).

36 Solopova, A. *et al.* Riboflavin Biosynthesis and Overproduction by a Derivative of the Human Gut Commensal Bifidobacterium longum subsp. infantis ATCC 15697. *Front Microbiol* 11, 573335, doi:10.3389/fmicb.2020.573335 (2020).

37 Sharma, V. *et al.* B-Vitamin Sharing Promotes Stability of Gut Microbial Communities. *Front Microbiol* 10, 1485, doi:10.3389/fmicb.2019.01485 (2019).

38 Kalliomaki, M. *et al.* Distinct patterns of neonatal gut microflora in infants in whom atopy was and was not developing. *J Allergy Clin Immunol* 107, 129-134, doi:10.1067/mai.2001.111237 (2001).

39 Fujimura, K. E. *et al.* Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nat Med* 22, 1187-1191, doi:10.1038/nm.4176 (2016).

40 Kalliomaki, M., Collado, M. C., Salminen, S. & Isolauri, E. Early differences in fecal microbiota composition in children may predict overweight. *Am J Clin Nutr* 87, 534-538, doi:10.1093/ajcn/87.3.534 (2008).

41 Taft, D. H. *et al.* Bifidobacterial Dominance of the Gut in Early Life and Acquisition of Antimicrobial Resistance. *mSphere* 3, doi:10.1128/mSphere.00441-18 (2018).

42 Szajewska, H. & Hojsak, I. Health benefits of Lactobacillus rhamnosus GG and Bifidobacterium animalis subspecies lactis BB-12 in children. *Postgrad Med* 132, 441-451, doi:10.1080/00325481.2020.1731214 (2020).

43 Wei, X. *et al.* Fructose uptake in Bifidobacterium longum NCC2705 is mediated by an ATP-binding cassette transporter. *J Biol Chem* 287, 357-367, doi:10.1074/jbc.M111.266213 (2012).

44 Liu, D. *et al.* Proteomics analysis of Bifidobacterium longum NCC2705 growing on glucose, fructose, mannose, xylose, ribose, and galactose. *Proteomics* 11, 2628-2638, doi:10.1002/pmic.201100035 (2011).

45 Palframan, R. J., Gibson, G. R. & Rastall, R. A. Carbohydrate preferences of Bifidobacterium species isolated from the human gut. *Curr Issues Intest Microbiol* 4, 71-75 (2003).

46 Pokusaeva, K., Fitzgerald, G. F. & van Sinderen, D. Carbohydrate metabolism in Bifidobacteria. *Genes Nutr* 6, 285-306, doi:10.1007/s12263-010-0206-6 (2011).

47 Cordeiro, R. L. *et al.* N-glycan Utilization by Bifidobacterium Gut Symbionts Involves a Specialist beta-Mannosidase. *J Mol Biol* 431, 732-747, doi:10.1016/j.jmb.2018.12.017 (2019).

48 Kullin, B., Abratt, V. R. & Reid, S. J. A functional analysis of the Bifidobacterium longum cscA and scrP genes in sucrose utilization. *Appl Microbiol Biotechnol* 72, 975-981, doi:10.1007/s00253-006-0358-x (2006).

49 Hinz, S. W., Pastink, M. I., van den Broek, L. A., Vincken, J. P. & Voragen, A. G. Bifidobacterium longum endogalactanase liberates galactotriose from type I galactans. *Appl Environ Microbiol* 71, 5501-5510, doi:10.1128/AEM.71.9.5501-5510.2005 (2005).

50 Savard, P. & Roy, D. Determination of Differentially Expressed Genes Involved in Arabinoxylan Degradation by Bifidobacterium longum NCC2705 Using Real-Time RT-PCR. *Probiotics and Antimicrobial Proteins* 1, 121, doi:10.1007/s12602-009-9015-x (2009).

51 Gonzalez-Rodriguez, I. *et al.* Catabolism of glucose and lactose in Bifidobacterium animalis subsp. lactis, studied by 13C Nuclear Magnetic Resonance. *Appl Environ Microbiol* 79, 7628-7638, doi:10.1128/AEM.02529-13 (2013).

52 Yuan, J. *et al.* Analysis of host-inducing proteome changes in bifidobacterium longum NCC2705 grown in Vivo. *J Proteome Res* 7, 375-385, doi:10.1021/pr0704940 (2008).

53 Price, C. E., Reid, S. J., Driessen, A. J. & Abratt, V. R. The Bifidobacterium longum NCIMB 702259T ctr gene codes for a novel cholate transporter. *Appl Environ Microbiol* 72, 923-926, doi:10.1128/AEM.72.1.923-926.2006 (2006).

54 Gueimonde, M., Garrigues, C., van Sinderen, D., de los Reyes-Gavilan, C. G. & Margolles, A. Bile-inducible efflux transporter from Bifidobacterium longum NCC2705, conferring bile resistance. *Appl Environ Microbiol* 75, 3153-3160, doi:10.1128/AEM.00172-09 (2009).

55 Wei, X. *et al.* Proteomic analysis of the interaction of Bifidobacterium longum NCC2705 with the intestine cells Caco-2 and identification of plasminogen receptors. *J Proteomics* 108, 89-98, doi:10.1016/j.jprot.2014.04.038 (2014).

56 Nishiyama, K. *et al.* Extracellular Vesicles Produced by Bifidobacterium longum Export Mucin-Binding Proteins. *Appl Environ Microbiol* 86, doi:10.1128/AEM.01464-20 (2020).

57 Suzuki, K. *et al.* Adhesion properties of a putative polymorphic fimbrial subunit protein from Bifidobacterium longum subsp. longum. *Biosci Microbiota Food Health* 35, 19-27, doi:10.12938/bmfh.2015-015 (2016).

58 Oberg, T. S. *et al.* Intrinsic and inducible resistance to hydrogen peroxide in Bifidobacterium species. *J Ind Microbiol Biotechnol* 38, 1947-1953, doi:10.1007/s10295-011-0983-y (2011).

59 Zuo, F. *et al.* Transcriptomic analysis of Bifidobacterium longum subsp. longum BBMN68 in response to oxidative shock. *Sci Rep* 8, 17085, doi:10.1038/s41598-018-35286-7 (2018).

60 Zuo, F. *et al.* Homologous overexpression of alkyl hydroperoxide reductase subunit C (ahpC) protects Bifidobacterium longum strain NCC2705 from oxidative stress. *Res Microbiol* 165, 581-589, doi:10.1016/j.resmic.2014.05.040 (2014).

61 Ivanov, D. *et al.* A serpin from the gut bacterium Bifidobacterium longum inhibits eukaryotic elastase-like serine proteases. *J Biol Chem* 281, 17246-17252, doi:10.1074/jbc.M601678200 (2006).

62 Mundi, A., Delcenserie, V., Amiri-Jami, M., Moorhead, S. & Griffiths, M. W. Cell-free preparations of Lactobacillus acidophilus strain La-5 and Bifidobacterium longum strain NCC2705 affect virulence gene expression in Campylobacter jejuni. *J Food Prot* 76, 1740-1746, doi:10.4315/0362-028X.JFP-13-084 (2013).

63 Riedel, C. U. *et al.* Anti-inflammatory effects of bifidobacteria by inhibition of LPS-induced NF-kappaB activation. *World J Gastroenterol* 12, 3729-3735 (2006).

64 McCarville, J. L. *et al.* A commensal Bifidobacterium longum strain improves gluten-related immunopathology in mice through expression of a serine protease inhibitor. *Appl Environ Microbiol*, doi:10.1128/AEM.01323-17 (2017).

65 Burg, N. D. & Pillinger, M. H. The neutrophil: function and regulation in innate and humoral immunity. *Clin Immunol* 99, 7-17, doi:10.1006/clim.2001.5007 (2001).

66 Menard, O., Butel, M. J., Gaboriau-Routhiau, V. & Waligora-Dupriet, A. J. Gnotobiotic mouse immune response induced by Bifidobacterium sp. strains isolated from infants. *Appl Environ Microbiol* 74, 660-666, doi:10.1128/AEM.01261-07 (2008).

67 Menard, O. *et al.* Characterization of immunostimulatory CpG-rich sequences from different Bifidobacterium species. *Appl Environ Microbiol* 76, 2846-2855, doi:10.1128/AEM.01714-09 (2010).

68 Otten, B. *et al.* Safety of bifidobacterium longum NCC 2705 and production of its serpin in patients with celiac disease and non-celiac gluten sensitivity. *United european gastroenterology journal* 8, 118-119, doi:10.1177/2050640620927344 (2020).

69 FAO/WHO. Guidelines for the Evaluation of Probiotics in Food. (Working Group Report on Drafting Guidelines for the Evaluation of Probiotics in Food, 2002).

70 Hill, C. *et al.* Expert consensus document. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nat Rev Gastroenterol Hepatol* 11, 506-514, doi:10.1038/nrgastro.2014.66 (2014).

71 Fenster, K. *et al.* The Production and Delivery of Probiotics: A Review of a Practical Approach. *Microorganisms* 7, doi:10.3390/microorganisms7030083 (2019).

72 Fiocco, D. *et al.* How probiotics face food stress: They get by with a little help. *Critical Reviews in Food Science and Nutrition* 60, 1552-1580, doi:10.1080/10408398.2019.1580673 (2020).

73 Corcoran, B. M., Stanton, C., Fitzgerald, G. & Ross, R. P. Life under stress: the probiotic stress response and how it may be manipulated. *Curr Pharm Des* 14, 1382-1399, doi:10.2174/138161208784480225 (2008).

74 Mills, S., Stanton, C., Fitzgerald, G. F. & Ross, R. P. Enhancing the stress responses of probiotics for a lifestyle from gut to product and back again. *Microb Cell Fact* 10 Suppl 1, S19, doi:10.1186/1475-2859-10-S1-S19 (2011).

75 Wei, Y., Gao, J., Liu, D., Li, Y. & Liu, W. Adaptational changes in physiological and transcriptional responses of Bifidobacterium longum involved in acid stress resistance after successive batch cultures. *Microb Cell Fact* 18, 156, doi:10.1186/s12934-019-1206-x (2019).

76 Sánchez, B. *et al.* Low-pH Adaptation and the Acid Tolerance Response of Bifidobacterium longum Biotype longum. *Applied and Environmental Microbiology* 73, 6450-6459, doi:10.1128/aem.00886-07 (2007).

77 Ruiz, L., Sanchez, B., Ruas-Madiedo, P., de Los Reyes-Gavilan, C. G. & Margolles, A. Cell envelope changes in Bifidobacterium animalis ssp. lactis as a response to bile. *FEMS Microbiol Lett* 274, 316-322, doi:10.1111/j.1574-6968.2007.00854.x (2007).

78 Schopping, M., Gaspar, P., Neves, A. R., Franzen, C. J. & Zeidan, A. A. Identifying the essential nutritional requirements of the probiotic bacteria Bifidobacterium animalis and Bifidobacterium longum through genome-scale modeling. *NPJ Syst Biol Appl* 7, 47, doi:10.1038/s41540-021-00207-4 (2021).

79 Ummadi, M. S. & Curic-Bawden, M. in *Protein hydrolysates in biotechnology* 91-114 (Springer, 2008).

80 Milić, T. V., Rakin, M. & Šiler-Marinković, S. Utilization of baker's yeast (Saccharomyces cerevisiae) for the production of yeast extract: Effects of different enzymatic treatments on solid, protein and carbohydrate recovery. . *Journal of the Serbian Chemical Society* 72, 451-457, doi:https://doi.org/10.2298/JSC0705451V (2007).

81 Verduyn, C., Suksomcheep, A. & Suphantharika, M. Effect of high pressure homogenization and papain on the preparation of autolysed yeast extract. *World journal of microbiology & biotechnology* 1999 v.15 no.1, pp. 57-63, doi:10.1023/a:1008818511497 (1999).

82 Gray, V. L., Muller, C. T., Watkins, I. D. & Lloyd, D. Peptones from diverse sources: pivotal determinants of bacterial growth dynamics. *J Appl Microbiol* 104, 554-565, doi:10.1111/j.1365-2672.2007.03577.x (2008).

83 Doleyres, Y. & Lacroix, C. Technologies with free and immobilised cells for probiotic bifidobacteria production and protection. *International Dairy Journal* 15, 973-988, doi:https://doi.org/10.1016/j.idairyj.2004.11.014 (2005).

84 Duboux, S., Van Wijchen, M. & Kleerebezem, M. The Possible Link Between Manufacturing and Probiotic Efficacy; a Molecular Point of View on Bifidobacterium. *Front Microbiol* 12, 812536, doi:10.3389/fmicb.2021.812536 (2021).

85 Joint, F. WHO working group report on drafting guidelines for the evaluation of probiotics in food. *London, Ontario, Canada* 30 (2002).

86 Dronkers, T. M. G., Ouwehand, A. C. & Rijkers, G. T. Global analysis of clinical trials with probiotics. *Heliyon* 6, e04467, doi:10.1016/j.heliyon.2020.e04467 (2020).

87 Lebeer, S., Vanderleyden, J. & De Keersmaecker, S. C. Genes and molecules of lactobacilli supporting probiotic action. *Microbiol Mol Biol Rev* 72, 728-764, Table of Contents, doi:10.1128/MMBR.00017-08 (2008).

88 Lee, I. C., Tomita, S., Kleerebezem, M. & Bron, P. A. The quest for probiotic effector molecules--unraveling strain specificity at the molecular level. *Pharmacol Res* 69, 61-74, doi:10.1016/j.phrs.2012.09.010 (2013).

89 Remus, D., Kleerebezem, M. & Bron, P. A. *Molecular analysis of candidate Probiotic Effector Molecules of Lactobacillus plantarum* PhD manuscript thesis, Wageningen University, the Netherlands, (2012).

90 Lebeer, S. *et al.* Identification of probiotic effector molecules: present state and future perspectives. *Curr Opin Biotechnol* 49, 217-223, doi:10.1016/j.copbio.2017.10.007 (2018).

91 Osborn, D. A. & Sinn, J. K. Probiotics in infants for prevention of allergic disease and food hypersensitivity. *Cochrane Database Syst Rev*, CD006475, doi:10.1002/14651858.CD006475.pub2 (2007).

92 Johnston, B. C., Goldenberg, J. Z., Vandvik, P. O., Sun, X. & Guyatt, G. H. Probiotics for the prevention of pediatric antibiotic-associated diarrhea. *Cochrane Database Syst Rev*, CD004827, doi:10.1002/14651858.CD004827.pub3 (2011).

93 Guo, Q., Goldenberg, J. Z., Humphrey, C., El Dib, R. & Johnston, B. C. Probiotics for the prevention of pediatric antibiotic-associated diarrhea. *Cochrane Database Syst Rev* 4, CD004827, doi:10.1002/14651858.CD004827.pub5 (2019).

94 Forssten, S. D., Laitila, A., Maukonen, J. & Ouwehand, A. C. Probiotic triangle of success; strain production, clinical studies and product development. *FEMS Microbiology Letters* 367, doi:10.1093/femsle/fnaa167 (2020).

95 Veiga, P., Suez, J., Derrien, M. & Elinav, E. Moving from probiotics to precision probiotics. *Nat Microbiol* 5, 878-880, doi:10.1038/s41564-020-0721-1 (2020).

96 Segers, M. E. & Lebeer, S. Towards a better understanding of Lactobacillus rhamnosus GG--host interactions. *Microb Cell Fact* 13 Suppl 1, S7, doi:10.1186/1475-2859-13-S1-S7 (2014).

97 Kalliomaki, M. *et al.* Probiotics in primary prevention of atopic disease: a randomised placebo-controlled trial. *Lancet* 357, 1076-1079, doi:10.1016/S0140-6736(00)04259-8 (2001).

98 Kalliomaki, M., Salminen, S., Poussa, T. & Isolauri, E. Probiotics during the first 7 years of life: a cumulative risk reduction of eczema in a randomized, placebo-controlled trial. *J Allergy Clin Immunol* 119, 1019-1021, doi:10.1016/j.jaci.2006.12.608 (2007).

99 Kopp, M. V., Hennemuth, I., Heinzmann, A. & Urbanek, R. Randomized, double-blind, placebo-controlled trial of probiotics for primary prevention: no clinical effects of Lactobacillus GG supplementation. *Pediatrics* 121, e850-856, doi:10.1542/peds.2007-1492 (2008).

100 Tripathi, P. *et al.* Deciphering the nanometer-scale organization and assembly of Lactobacillus rhamnosus GG pili using atomic force microscopy. *Langmuir* 28, 2211-2216, doi:10.1021/la203834d (2012).

101 Brussow, H. Probiotics and prebiotics in clinical tests: an update. *F1000Res* 8, doi:10.12688/f1000research.19043.1 (2019).

102 Sanders, M. E. *et al.* Effects of genetic, processing, or product formulation changes on efficacy and safety of probiotics. *Ann N Y Acad Sci* 1309, 1-18, doi:10.1111/nyas.12363 (2014).

103 Gaucher, F. *et al.* Review: Adaptation of Beneficial Propionibacteria, Lactobacilli, and Bifidobacteria Improves Tolerance Toward Technological and Digestive Stresses. *Front Microbiol* 10, 841, doi:10.3389/fmicb.2019.00841 (2019).

104 Pique, N., Berlanga, M. & Minana-Galbis, D. Health Benefits of Heat-Killed (Tyndallized) Probiotics: An Overview. *Int J Mol Sci* 20, doi:10.3390/ijms20102534 (2019).

105 Vargas Garcia, C. E. *et al.* Piliation of Lactobacillus rhamnosus GG promotes adhesion, phagocytosis, and cytokine modulation in macrophages. *Appl Environ Microbiol* 81, 2050-2062, doi:10.1128/AEM.03949-14 (2015).

106 Lebeer, S. *et al.* Functional analysis of Lactobacillus rhamnosus GG pili in relation to adhesion and immunomodulatory interactions with intestinal epithelial cells. *Appl Environ Microbiol* 78, 185-193, doi:10.1128/AEM.06192-11 (2012).

107 Ardita, C. S. *et al.* Epithelial adhesion mediated by pilin SpaC is required for Lactobacillus rhamnosus GG-induced cellular responses. *Appl Environ Microbiol* 80, 5068-5077, doi:10.1128/AEM.01039-14 (2014).

108 Yan, F. *et al.* A Lactobacillus rhamnosus GG-derived soluble protein, p40, stimulates ligand release from intestinal epithelial cells to transactivate epidermal growth factor receptor. *J Biol Chem* 288, 30742-30751, doi:10.1074/jbc.M113.492397 (2013).

109 Debabov, D. V., Kiriukhin, M. Y. & Neuhaus, F. C. Biosynthesis of lipoteichoic acid in Lactobacillus rhamnosus: role of DltD in D-alanylation. *J Bacteriol* 182, 2855-2864, doi:10.1128/JB.182.10.2855-2864.2000 (2000).

110 Debabov, D. V. *et al.* The D-Alanyl carrier protein in Lactobacillus casei: cloning, sequencing, and expression of dltC. *J Bacteriol* 178, 3869-3876, doi:10.1128/jb.178.13.3869-3876.1996 (1996).

111 Claes, I. J. *et al.* Lipoteichoic acid is an important microbe-associated molecular pattern of Lactobacillus rhamnosus GG. *Microb Cell Fact* 11, 161, doi:10.1186/1475-2859-11-161 (2012).

112 Koprivnjak, T. *et al.* Cation-induced transcriptional regulation of the dlt operon of Staphylococcus aureus. *J Bacteriol* 188, 3622-3630, doi:10.1128/JB.188.10.3622-3630.2006 (2006).

113 Revilla-Guarinos, A. *et al.* Characterization of a regulatory network of peptide antibiotic detoxification modules in Lactobacillus casei BL23. *Appl Environ Microbiol* 79, 3160-3170, doi:10.1128/AEM.00178-13 (2013).

114 Lebeer, S. *et al.* Identification of a Gene Cluster for the Biosynthesis of a Long, Galactose-Rich Exopolysaccharide in Lactobacillus rhamnosus GG and Functional Analysis of the Priming Glycosyltransferase. *Appl Environ Microbiol* 75, 3554-3563, doi:10.1128/AEM.02919-08 (2009).

115 Nguyen, P. T. *et al.* Response of Lactobacillus plantarum VAL6 to challenges of pH and sodium chloride stresses. *Sci Rep* 11, 1301, doi:10.1038/s41598-020-80634-1 (2021).

116 Kiekens, S. *et al.* Impact of spray-drying on the pili of Lactobacillus rhamnosus GG. *Microb Biotechnol* 12, 849-855, doi:10.1111/1751-7915.13426 (2019).

117 du Toit, E., Vesterlund, S., Gueimonde, M. & Salminen, S. Assessment of the effect of stress-tolerance acquisition on some basic characteristics of specific probiotics. *Int J Food Microbiol* 165, 51-56, doi:10.1016/j.ijfoodmicro.2013.04.022 (2013).

118 Sybesma, W., Molenaar, D., van, I. W., Venema, K. & Kort, R. Genome instability in Lactobacillus rhamnosus GG. *Appl Environ Microbiol* 79, 2233-2239, doi:10.1128/AEM.03566-12 (2013).

119 Grzeskowiak, L., Isolauri, E., Salminen, S. & Gueimonde, M. Manufacturing process influences properties of probiotic bacteria. *Br J Nutr* 105, 887-894, doi:10.1017/S0007114510004496 (2011).

120 van Baarlen, P. *et al.* Differential NF-kappaB pathways induction by Lactobacillus plantarum in the duodenum of healthy humans correlating with immune tolerance. *Proc Natl Acad Sci U S A* 106, 2371-2376, doi:10.1073/pnas.0809919106 (2009).

121 Bron, P. A., Grangette, C., Mercenier, A., de Vos, W. M. & Kleerebezem, M. Identification of Lactobacillus plantarum genes that are induced in the gastrointestinal tract of mice. *J Bacteriol* 186, 5721-5729, doi:10.1128/JB.186.17.5721-5729.2004 (2004).

122 Marco, M. L., Bongers, R. S., de Vos, W. M. & Kleerebezem, M. Spatial and temporal expression of Lactobacillus plantarum genes in the gastrointestinal tracts of mice. *Appl Environ Microbiol* 73, 124-132, doi:10.1128/AEM.01475-06 (2007).

123 Marco, M. L. *et al.* Convergence in probiotic Lactobacillus gut-adaptive responses in humans and mice. *ISME J* 4, 1481-1484, doi:10.1038/ismej.2010.61 (2010).

124 Laconelli, C. *et al.* Drying process strongly affects probiotics viability and functionalities. *J Biotechnol* 214, 17-26, doi:10.1016/j.jbiotec.2015.08.022 (2015).

125 Burns, P. *et al.* Spray-drying process preserves the protective capacity of a breast milk-derived Bifidobacterium lactis strain on acute and chronic colitis in mice. *Sci Rep* 7, 43211, doi:10.1038/srep43211 (2017).

126 Charnchai, P., Jantama, S. S., Prasitpuriprecha, C., Kanchanatawee, S. & Jantama, K. Effects of the Food Manufacturing Chain on the Viability and Functionality of Bifidobacterium animalis through Simulated Gastrointestinal Conditions. *PLoS One* 11, e0157958, doi:10.1371/journal.pone.0157958 (2016).

127 Westermann, C., Gleinser, M., Corr, S. C. & Riedel, C. U. A Critical Evaluation of Bifidobacterial Adhesion to the Host Tissue. *Front Microbiol* 7, 1220-1220, doi:10.3389/fmicb.2016.01220 (2016).

128 O'Connell Motherway, M. *et al.* Functional genome analysis of Bifidobacterium breve UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor. *Proceedings of the National Academy of Sciences of the United States of America* 108, 11217-11222, doi:10.1073/pnas.1105380108 (2011).

129 Foroni, E. *et al.* in *Microbial Cell Factories.* S16 (BioMed Central).

130 Duranti, S. *et al.* Genomic Characterization and Transcriptional Studies of the Starch-Utilizing Strain Bifidobacterium adolescentis 22L. *Applied and Environmental Microbiology* 80, 6080-6090, doi:10.1128/aem.01993-14 (2014).

131 Turroni, F. *et al.* Role of sortase-dependent pili of Bifidobacterium bifidum PRL2010 in modulating bacterium–host interactions. *Proceedings of the National Academy of Sciences* 110, 11151-11156, doi:10.1073/pnas.1303897110 (2013).
132 Pu, M., Duriez, P., Arazi, M. & Rowe-Magnus, D. A. A conserved tad pilus promotes Vibrio vulnificus oyster colonization. *Environmental microbiology* 20, 828-841 (2018).
133 Westermann, C. *et al.* Exploring the genome sequence of Bifidobacterium bifidum S17 for potential players in host-microbe interactions. *Symbiosis* 58, 191-200, doi:10.1007/s13199-012-0205-z (2012).
134 O'Connell Motherway, M. *et al.* A Bifidobacterial pilus-associated protein promotes colonic epithelial proliferation. *Mol Microbiol* 111, 287-301, doi:10.1111/mmi.14155 (2019).
135 Serafini, F. *et al.* Kefir fermented milk and kefiran promote growth of Bifidobacterium bifidum PRL2010 and modulate its gene expression. *International Journal of Food Microbiology* 178, 50-59, doi:https://doi.org/10.1016/j.ijfoodmicro.2014.02.024 (2014).
136 Turroni, F. *et al.* Expression of sortase-dependent pili of Bifidobacterium bifidum PRL2010 in response to environmental gut conditions. *FEMS Microbiology Letters* 357, 23-33, doi:10.1111/1574-6968.12509 (2014).
137 Turroni, F. *et al.* Characterization of the serpin-encoding gene of Bifidobacterium breve 210B. *Appl Environ Microbiol* 76, 3206-3219, doi:10.1128/AEM.02938-09 (2010).
138 Buhner, S. *et al.* Protease signaling through protease activated receptor 1 mediate nerve activation by mucosal supernatants from irritable bowel syndrome but not from ulcerative colitis patients. *PLoS One* 13, e0193943, doi:10.1371/journal.pone.0193943 (2018).
139 Alvarez-Martin, P. *et al.* A two-component regulatory system controls autoregulated serpin expression in Bifidobacterium breve UCC2003. *Appl Environ Microbiol* 78, 7032-7041, doi:10.1128/AEM.01776-12 (2012).
140 Duboux, S. *et al.* Carbohydrate-controlled serine protease inhibitor (serpin) production in Bifidobacterium longum subsp. longum. *Sci Rep* 11, 7236, doi:10.1038/s41598-021-86740-y (2021).
141 Leivers, S. *et al.* Structure of the high molecular weight exopolysaccharide produced by Bifidobacterium animalis subsp. lactis IPLA-R1 and sequence analysis of its putative eps cluster. *Carbohydrate research* 346, 2710-2717 (2011).
142 Hidalgo-Cantabrana, C. *et al.* Immune modulation capability of exopolysaccharides synthesised by lactic acid bacteria and bifidobacteria. *Probiotics and Antimicrobial Proteins* 4, 227-237 (2012).
143 Altmann, F. *et al.* Genome analysis and characterisation of the exopolysaccharide produced by Bifidobacterium longum subsp. longum 35624™. *PloS one* 11, e0162983 (2016).
144 Schiavi, E. *et al.* The Surface-Associated Exopolysaccharide of Bifidobacterium longum 35624 Plays an Essential Role in Dampening Host Proinflammatory Responses and Repressing Local TH17 Responses. *Applied and Environmental Microbiology* 82, 7185-7196, doi:10.1128/aem.02238-16 (2016).
145 Castro-Bravo, N., Hidalgo-Cantabrana, C., Rodriguez-Carvajal, M. A., Ruas-Madiedo, P. & Margolles, A. Gene Replacement and Fluorescent Labeling to Study the Functional Role of Exopolysaccharides in Bifidobacterium animalis subsp. lactis. *Frontiers in microbiology* 8, 1405-1405, doi:10.3389/fmicb.2017.01405 (2017).
146 Xu, R., Shen, Q., Ding, X., Gao, W. & Li, P. Chemical characterization and antioxidant activity of an exopolysaccharide fraction isolated from Bifidobacterium animalis RH. *European Food Research and Technology = Zeitschrift für Lebensmittel-Untersuchung und -Forschung. A* 232, 231-240, doi:10.1007/s00217-010-1382-8 (2011).
147 Amiri, S., Rezaei Mokarram, R., Sowti Khiabani, M., Rezazadeh Bari, M. & Alizadeh Khaledabad, M. Exopolysaccharides production by Lactobacillus acidophilus LA5 and Bifidobacterium animalis subsp. lactis BB12: Optimization of fermentation variables and characterization of structure and bioactivities. *International Journal of Biological Macromolecules* 123, 752-765, doi:https://doi.org/10.1016/j.ijbiomac.2018.11.084 (2019).

148 Hidalgo-Cantabrana, C. *et al.* Genomic Overview and Biological Functions of Exopolysaccharide Biosynthesis in Bifidobacterium spp. *Applied and Environmental Microbiology* 80, 9-18, doi:10.1128/aem.02977-13 (2014).
149 Wu, M.-H. *et al.* Exopolysaccharide activities from probiotic bifidobacterium: Immunomodulatory effects (on J774A.1 macrophages) and antimicrobial properties. *International Journal of Food Microbiology* 144, 104-110, doi:https://doi.org/10.1016/j.ijfoodmicro.2010.09.003 (2010).
150 López, P. *et al.* Exopolysaccharide-producing Bifidobacterium strains elicit different in vitro responses upon interaction with human cells. *Food Research International* 46, 99-107, doi:https://doi.org/10.1016/j.foodres.2011.11.020 (2012).
151 Xu, R., Shen, Q., Wu, R. & Li, P. Structural analysis and mucosal immune regulation of exopolysaccharide fraction from Bifidobacterium animalis RH. *Food and Agricultural Immunology* 28, 1226-1241, doi:10.1080/09540105.2017.1333578 (2017).
152 Hidalgo-Cantabrana, C. *et al.* Effect of a ropy exopolysaccharide-producing Bifidobacterium animalis subsp. lactis strain orally administered on DSS-induced colitis mice model. *Frontiers in microbiology* 7, 868 (2016).
153 Hidalgo-Cantabrana, C. *et al.* Insights into the ropy phenotype of the exopolysaccharide-producing strain Bifidobacterium animalis subsp. lactis A1dOxR. *Appl Environ Microbiol* 79, 3870-3874, doi:10.1128/AEM.00633-13 (2013).
154 Hidalgo-Cantabrana, C. *et al.* A single mutation in the gene responsible for the mucoid phenotype of Bifidobacterium animalis subsp. lactis confers surface and functional characteristics. *Appl Environ Microbiol* 81, 7960-7968, doi:10.1128/AEM.02095-15 (2015).
155 Audy, J., Labrie, S., Roy, D. & LaPointe, G. Sugar source modulates exopolysaccharide biosynthesis in Bifidobacterium longum subsp. longum CRC 002. *Microbiology* 156, 653-664 (2010).
156 Roberts, C. M. *et al.* Exopolysaccharide production by Bifidobacterium longum BB-79. *Journal of Applied Bacteriology* 78, 463-468, doi:10.1111/j.1365-2672.1995.tb03085.x (1995).
157 Amiri, S., Rezaei Mokarram, R., Sowti Khiabani, M., Rezazadeh Bari, M. & Alizadeh Khaledabad, M. Exopolysaccharides production by Lactobacillus acidophilus LA5 and Bifidobacterium animalis subsp. lactis BB12: Optimization of fermentation variables and characterization of structure and bioactivities. *Int J Biol Macromol* 123, 752-765, doi:10.1016/j.ijbiomac.2018.11.084 (2019).
158 Ninomiya, K. *et al.* Effect of CO2 concentration on the growth and exopolysaccharide production of Bifidobacterium longum cultivated under anaerobic conditions. *Journal of Bioscience and Bioengineering* 107, 535-537, doi:https://doi.org/10.1016/j.jbiosc.2008.12.015 (2009).
159 BERECKA, M. P., WAŚKO, A., Szwajgier, D. & Choma, A. Bifidogenic and antioxidant activity of exopolysaccharides produced by Lactobacillus rhamnosus E/N cultivated on different carbon sources. *Polish journal of microbiology* 62, 181-189 (2013).
160 Gonzalez-Rodriguez, I. *et al.* Role of extracellular transaldolase from Bifidobacterium bifidum in mucin adhesion and aggregation. *Appl Environ Microbiol* 78, 3992-3998, doi:10.1128/AEM.08024-11 (2012).
161 Candela, M. *et al.* Binding of human plasminogen to Bifidobacterium. *Journal of bacteriology* 189, 5929-5936 (2007).
162 Gleinser, M., Grimm, V., Zhurina, D., Yuan, J. & Riedel, C. U. Improved adhesive properties of recombinant bifidobacteria expressing the Bifidobacterium bifidum-specific lipoprotein BopA. *Microbial cell factories* 11, 80 (2012).
163 Jeffery, C. J. Moonlighting proteins. *Trends Biochem Sci* 24, 8-11, doi:10.1016/s0968-0004(98)01335-8 (1999).
164 Toyofuku, M. *et al.* Prophage-triggered membrane vesicle formation through peptidoglycan damage in Bacillus subtilis. *Nat Commun* 8, 481, doi:10.1038/s41467-017-00492-w (2017).

165 Candela, M. *et al.* DnaK from Bifidobacterium animalis subsp. lactis is a surface-exposed human plasminogen receptor upregulated in response to bile salts. *Microbiology* 156, 1609-1618 (2010).

166 González-Rodríguez, I. *et al.* Role of Extracellular Transaldolase from Bifidobacterium bifidum in Mucin Adhesion and Aggregation. *Applied and Environmental Microbiology* 78, 3992-3998, doi:10.1128/aem.08024-11 (2012).

167 Hidalgo-Cantabrana, C. *et al.* Insights into the Ropy Phenotype of the Exopolysaccharide-Producing Strain Bifidobacterium animalis subsp. lactis A1dOxR. *Applied and Environmental Microbiology* 79, 3870-3874, doi:10.1128/aem.00633-13 (2013).

168 Turroni, F. *et al.* Bifidobacteria and the infant gut: an example of co-evolution and natural selection. *Cell Mol Life Sci* 75, 103-118, doi:10.1007/s00018-017-2672-0 (2018).

169 Ayechu-Muruzabal, V. *et al.* Diversity of Human Milk Oligosaccharides and Effects on Early Life Immune Development. *Front Pediatr* 6, 239, doi:10.3389/fped.2018.00239 (2018).

170 Jakobsen, L. M. A., Sundekilde, U. K., Andersen, H. J., Nielsen, D. S. & Bertram, H. C. Lactose and Bovine Milk Oligosaccharides Synergistically Stimulate B. longum subsp. longum Growth in a Simplified Model of the Infant Gut Microbiome. *J Proteome Res*, doi:10.1021/acs.jproteome.9b00211 (2019).

171 Meli, F. *et al.* Growth and safety evaluation of infant formulae containing oligosaccharides derived from bovine milk: a randomized, double-blind, noninferiority trial. *BMC Pediatr* 14, 306, doi:10.1186/s12887-014-0306-3 (2014).

172 Tandon, D. *et al.* A prospective randomized, double-blind, placebo-controlled, dose-response relationship study to investigate efficacy of fructo-oligosaccharides (FOS) on human gut microflora. *Sci Rep* 9, 5473, doi:10.1038/s41598-019-41837-3 (2019).

173 O'Leary, N. A. *et al.* Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* 44, D733-745, doi:10.1093/nar/gkv1189 (2016).

174 Mitchell, A. L. *et al.* InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res* 47, D351-D360, doi:10.1093/nar/gky1100 (2019).

175 Almagro Armenteros, J. J. *et al.* SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* 37, 420-423, doi:10.1038/s41587-019-0036-z (2019).

176 Deutscher, J. The mechanisms of carbon catabolite repression in bacteria. *Curr Opin Microbiol* 11, 87-93, doi:10.1016/j.mib.2008.02.007 (2008).

177 Caescu, C. I., Vidal, O., Krzewinski, F., Artenie, V. & Bouquelet, S. Bifidobacterium longum requires a fructokinase (Frk; ATP:D-fructose 6-phosphotransferase, EC 2.7.1.4) for fructose catabolism. *J Bacteriol* 186, 6515-6525, doi:10.1128/JB.186.19.6515-6525.2004 (2004).

178 Trindade, M. I., Abratt, V. R. & Reid, S. J. Induction of sucrose utilization genes from Bifidobacterium lactis by sucrose and raffinose. *Appl Environ Microbiol* 69, 24-32, doi:10.1128/aem.69.1.24-32.2003 (2003).

179 Nessler, S. *et al.* HPr kinase/phosphorylase, the sensor enzyme of catabolite repression in Gram-positive bacteria: structural aspects of the enzyme and the complex with its protein substrate. *J Bacteriol* 185, 4003-4010, doi:10.1128/jb.185.14.4003-4010.2003 (2003).

180 Novichkov, P. S. *et al.* RegPrecise 3.0--a resource for genome-scale exploration of transcriptional regulation in bacteria. *BMC Genomics* 14, 745, doi:10.1186/1471-2164-14-745 (2013).

181 Lanigan, N. *et al.* Transcriptional control of central carbon metabolic flux in Bifidobacteria by two functionally similar, yet distinct LacI-type regulators. *Sci Rep* 9, 17851, doi:10.1038/s41598-019-54229-4 (2019).

182 Khoroshkin, M. S., Leyn, S. A., Van Sinderen, D. & Rodionov, D. A. Transcriptional Regulation of Carbohydrate Utilization Pathways in the Bifidobacterium Genus. *Front Microbiol* 7, 120, doi:10.3389/fmicb.2016.00120 (2016).

183 Bailey, T. L. *et al.* MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* 37, W202-208, doi:10.1093/nar/gkp335 (2009).

184 Ganzle, M. Enzymatic synthesis of galacto-oligosaccharides and other lactose derivatives (hetero-oligosaccharides) from lactose. *International Dairy Journal* 22, 116-122, doi:10.1016/j.idairyj.2011.06.010 (2001).
185 Martins, G. N., Ureta, M. M., Tymczyszyn, E. E., Castilho, P. C. & Gomez-Zavaglia, A. Technological Aspects of the Production of Fructo and Galacto-Oligosaccharides. Enzymatic Synthesis and Hydrolysis. *Front Nutr* 6, 78, doi:10.3389/fnut.2019.00078 (2019).
186 Ripio, M. T., Brehm, K., Lara, M., Suarez, M. & Vazquez-Boland, J. A. Glucose-1-phosphate utilization by Listeria monocytogenes is PrfA dependent and coordinately expressed with virulence factors. *J Bacteriol* 179, 7174-7180, doi:10.1128/jb.179.22.7174-7180.1997 (1997).
187 Sakurama, H. *et al.* Lacto-N-biosidase encoded by a novel gene of Bifidobacterium longum subspecies longum shows unique substrate specificity and requires a designated chaperone for its active expression. *J Biol Chem* 288, 25194-25206, doi:10.1074/jbc.M113.484733 (2013).
188 Kunz, C., Rudloff, S., Baier, W., Klein, N. & Strobel, S. Oligosaccharides in human milk: structural, functional, and metabolic aspects. *Annu Rev Nutr* 20, 699-722, doi:10.1146/annurev.nutr.20.1.699 (2000).
189 Taboada, B., Ciria, R., Martinez-Guerrero, C. E. & Merino, E. ProOpDB: Prokaryotic Operon DataBase. *Nucleic Acids Res* 40, D627-631, doi:10.1093/nar/gkr1020 (2012).
190 Tannock, G. W. A special fondness for lactobacilli. *Appl Environ Microbiol* 70, 3189-3194, doi:10.1128/AEM.70.6.3189-3194.2004 (2004).
191 Maus, J. E. & Ingham, S. C. Employment of stressful conditions during culture production to enhance subsequent cold- and acid-tolerance of bifidobacteria. *J Appl Microbiol* 95, 146-154, doi:10.1046/j.1365-2672.2003.01954.x (2003).
192 Guan, N. & Liu, L. Microbial response to acid stress: mechanisms and applications. *Appl Microbiol Biotechnol* 104, 51-65, doi:10.1007/s00253-019-10226-1 (2020).
193 Amor, K. B. *et al.* Multiparametric flow cytometry and cell sorting for the assessment of viable, injured, and dead bifidobacterium cells during bile salt stress. *Appl Environ Microbiol* 68, 5209-5216, doi:10.1128/aem.68.11.5209-5216.2002 (2002).
194 Kim, S. *et al.* PubChem in 2021: new data content and improved web interfaces. *Nucleic Acids Res* 49, D1388-D1395, doi:10.1093/nar/gkaa971 (2021).
195 Trcek, J., Mira, N. P. & Jarboe, L. R. Adaptation and tolerance of bacteria against acetic acid. *Appl Microbiol Biotechnol* 99, 6215-6229, doi:10.1007/s00253-015-6762-3 (2015).
196 Waddington, L., Cyr, T., Hefford, M., Hansen, L. T. & Kalmokoff, M. Understanding the acid tolerance response of bifidobacteria. *J Appl Microbiol* 108, 1408-1420, doi:10.1111/j.1365-2672.2009.04540.x (2010).
197 den Besten, H. M., Mataragas, M., Moezelaar, R., Abee, T. & Zwietering, M. H. Quantification of the effects of salt stress and physiological state on thermotolerance of Bacillus cereus ATCC 10987 and ATCC 14579. *Appl Environ Microbiol* 72, 5884-5894, doi:10.1128/AEM.00780-06 (2006).
198 Arzamasov, A. A., van Sinderen, D. & Rodionov, D. A. Comparative Genomics Reveals the Regulatory Complexity of Bifidobacterial Arabinose and Arabino-Oligosaccharide Utilization. *Front Microbiol* 9, 776, doi:10.3389/fmicb.2018.00776 (2018).
199 Serafini, F. *et al.* An efficient and reproducible method for transformation of genetically recalcitrant bifidobacteria. *FEMS Microbiol Lett* 333, 146-152, doi:10.1111/j.1574-6968.2012.02605.x (2012).
200 Fukiya, S., Sakanaka, M. & Yokota, A. in *The Bifidobacteria and Related Organisms* (eds Paola Mattarelli, Bruno Biavati, Wilhelm H. Holzapfel, & Brian J. B. Wood) 243-259 (Academic Press, 2018).
201 Sun, Z., Westermann, C., Yuan, J. & Riedel, C. U. Experimental determination and characterization of the gap promoter of Bifidobacterium bifidum S17. *Bioengineered* 5, 371-377, doi:10.4161/bioe.34423 (2014).

202 Ruiz, L. *et al.* Controlled gene expression in bifidobacteria by use of a bile-responsive element. *Appl Environ Microbiol* 78, 581-585, doi:10.1128/AEM.06611-11 (2012).
203 Klijn, A. *et al.* Construction of a reporter vector for the analysis of Bifidobacterium longum promoters. *Appl Environ Microbiol* 72, 7401-7405, doi:10.1128/AEM.01611-06 (2006).
204 Lu, C., Bentley, W. E. & Rao, G. A high-throughput approach to promoter study using green fluorescent protein. *Biotechnol Prog* 20, 1634-1640, doi:10.1021/bp049751l (2004).
205 Lu, C., Bentley, W. E. & Rao, G. Comparisons of oxidative stress response genes in aerobic Escherichia coli fermentations. *Biotechnol Bioeng* 83, 864-870, doi:10.1002/bit.10732 (2003).
206 Zhao, H., Thompson, R. B., Lockatell, V., Johnson, D. E. & Mobley, H. L. Use of green fluorescent protein to assess urease gene expression by uropathogenic Proteus mirabilis during experimental ascending urinary tract infection. *Infect Immun* 66, 330-335 (1998).
207 Garcia-Cayuela, T. *et al.* Fluorescent protein vectors for promoter analysis in lactic acid bacteria and Escherichia coli. *Appl Microbiol Biotechnol* 96, 171-181, doi:10.1007/s00253-012-4087-z (2012).
208 Ogaugwu, C. E., Cheng, Q., Fieck, A., Hurwitz, I. & Durvasula, R. Characterization of a Lactococcus lactis promoter for heterologous protein production. *Biotechnol Rep (Amst)* 17, 86-92, doi:10.1016/j.btre.2017.11.010 (2018).
209 Shimamura, S. *et al.* Relationship between oxygen sensitivity and oxygen metabolism of Bifidobacterium species. *J Dairy Sci* 75, 3296-3306, doi:10.3168/jds.S0022-0302(92)78105-3 (1992).
210 Montenegro-Rodriguez, C., Peiroten, A., Sanchez-Jimenez, A., Arques, J. L. & Landete, J. M. Analysis of gene expression of bifidobacteria using as the reporter an anaerobic fluorescent protein. *Biotechnol Lett* 37, 1405-1413, doi:10.1007/s10529-015-1802-8 (2015).
211 Landete, J. M. *et al.* Anaerobic green fluorescent protein as a marker of Bifidobacterium strains. *Int J Food Microbiol* 175, 6-13, doi:10.1016/j.ijfoodmicro.2014.01.008 (2014).
212 Grimm, V., Gleinser, M., Neu, C., Zhurina, D. & Riedel, C. U. Expression of fluorescent proteins in bifidobacteria for analysis of host-microbe interactions. *Appl Environ Microbiol* 80, 2842-2850, doi:10.1128/AEM.04261-13 (2014).
213 Chia, H. E., Zuo, T., Koropatkin, N. M., Marsh, E. N. G. & Biteen, J. S. Imaging living obligate anaerobic bacteria with bilin-binding fluorescent proteins. *Current Research in Microbial Sciences* 1, 1-6, doi:https://doi.org/10.1016/j.crmicr.2020.04.001 (2020).
214 Wagner, J. R., Brunzelle, J. S., Forest, K. T. & Vierstra, R. D. A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* 438, 325-331, doi:10.1038/nature04118 (2005).
215 Karunatilaka, K. *et al. Single-molecule imaging can be achieved in live obligate anaerobic bacteria*. Vol. 8590 PWB (SPIE, 2013).
216 Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* 177, 4121-4130, doi:10.1128/jb.177.14.4121-4130.1995 (1995).
217 Du, F. *et al.* Regulating the T7 RNA polymerase expression in E. coli BL21 (DE3) to provide more host options for recombinant protein production. *Microb Cell Fact* 20, 189, doi:10.1186/s12934-021-01680-6 (2021).
218 Solopova, A. *et al.* Bet-hedging during bacterial diauxic shift. *Proc Natl Acad Sci U S A* 111, 7427-7432, doi:10.1073/pnas.1320063111 (2014).
219 Tarracchini, C. *et al.* Phylogenomic disentangling of the Bifidobacterium longum subsp. infantis taxon. *Microb Genom* 7, doi:10.1099/mgen.0.000609 (2021).
220 Zabel, B. E. *et al.* Strain-specific strategies of 2'-fucosyllactose, 3-fucosyllactose, and difucosyllactose assimilation by Bifidobacterium longum subsp. infantis Bi-26 and ATCC 15697. *Sci Rep* 10, 15919, doi:10.1038/s41598-020-72792-z (2020).
221 LoCascio, R. G., Desai, P., Sela, D. A., Weimer, B. & Mills, D. A. Broad conservation of milk utilization genes in Bifidobacterium longum subsp. infantis as revealed by comparative

genomic hybridization. *Appl Environ Microbiol* 76, 7373-7381, doi:10.1128/AEM.00675-10 (2010).

222 Warner, J. B. & Lolkema, J. S. CcpA-dependent carbon catabolite repression in bacteria. *Microbiol Mol Biol Rev* 67, 475-490, doi:10.1128/MMBR.67.4.475-490.2003 (2003).

223 Gorke, B. & Stulke, J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat Rev Microbiol* 6, 613-624, doi:10.1038/nrmicro1932 (2008).

224 Ihssen, J. & Egli, T. Specific growth rate and not cell density controls the general stress response in Escherichia coli. *Microbiology (Reading)* 150, 1637-1648, doi:10.1099/mic.0.26849-0 (2004).

225 Berney, M., Weilenmann, H. U., Ihssen, J., Bassin, C. & Egli, T. Specific growth rate determines the sensitivity of Escherichia coli to thermal, UVA, and solar disinfection. *Appl Environ Microbiol* 72, 2586-2593, doi:10.1128/AEM.72.4.2586-2593.2006 (2006).

226 Lindqvist, R. & Barmark, G. Specific growth rate determines the sensitivity of Escherichia coli to lactic acid stress: implications for predictive microbiology. *Biomed Res Int* 2014, 471317, doi:10.1155/2014/471317 (2014).

227 Biselli, E., Schink, S. J. & Gerland, U. Slower growth of Escherichia coli leads to longer survival in carbon starvation due to a decrease in the maintenance rate. *Mol Syst Biol* 16, e9478, doi:10.15252/msb.20209478 (2020).

228 Degnan, B. A. & Macfarlane, G. T. Comparison of carbohydrate substrate preferences in eight species of bifidobacteria. *FEMS Microbiol Lett* 68, 151-156, doi:10.1016/0378-1097(91)90119-u (1991).

229 Roy, D. Media for the isolation and enumeration of bifidobacteria in dairy products. *Int J Food Microbiol* 69, 167-182, doi:10.1016/s0168-1605(01)00496-2 (2001).

230 Bruggeman, F. J., Planque, R., Molenaar, D. & Teusink, B. Searching for principles of microbial physiology. *FEMS Microbiol Rev* 44, 821-844, doi:10.1093/femsre/fuaa034 (2020).

231 Ercan, O., Wels, M., Smid, E. J. & Kleerebezem, M. Molecular and metabolic adaptations of Lactococcus lactis at near-zero growth rates. *Appl Environ Microbiol* 81, 320-331, doi:10.1128/AEM.02484-14 (2015).

232 Nugroho, A. D. W., Kleerebezem, M. & Bachmann, H. A Novel Method for Long-Term Analysis of Lactic Acid and Ammonium Production in Non-growing Lactococcus lactis Reveals Pre-culture and Strain Dependence. *Front Bioeng Biotechnol* 8, 580090, doi:10.3389/fbioe.2020.580090 (2020).

233 Margolles, A., Garcia, L., Sanchez, B., Gueimonde, M. & de los Reyes-Gavilan, C. G. Characterisation of a Bifidobacterium strain with acquired resistance to cholate--a preliminary study. *Int J Food Microbiol* 82, 191-198, doi:10.1016/s0168-1605(02)00261-1 (2003).

234 Tacconi, S. *et al.* Carbohydrate stress-related response in Bifidobacterium pseudolongum subsp. globosum. *Annals of Microbiology* 62, 1751-1756, doi:10.1007/s13213-012-0432-9 (2012).

235 Yang, X., Hang, X., Tan, J. & Yang, H. Differences in acid tolerance between Bifidobacterium breve BB8 and its acid-resistant derivative B. breve BB8dpH, revealed by RNA-sequencing and physiological analysis. *Anaerobe* 33, 76-84, doi:10.1016/j.anaerobe.2015.02.005 (2015).

236 Jin, J. *et al.* Effect of Pre-Stressing on the Acid-Stress Response in Bifidobacterium Revealed Using Proteomic and Physiological Approaches. *PLoS One* 10, e0117702, doi:10.1371/journal.pone.0117702 (2015).

237 Collado, M. C. & Sanz, Y. Induction of acid resistance in Bifidobacterium: a mechanism for improving desirable traits of potentially probiotic strains. *J Appl Microbiol* 103, 1147-1157, doi:10.1111/j.1365-2672.2007.03342.x (2007).

238 Rezzonico, E. *et al.* Global transcriptome analysis of the heat shock response of Bifidobacterium longum. *FEMS Microbiol Lett* 271, 136-145, doi:10.1111/j.1574-6968.2007.00704.x (2007).

239 An, H. *et al.* Integrated transcriptomic and proteomic analysis of the bile stress response in a centenarian-originated probiotic Bifidobacterium longum BBMN68. *Mol Cell Proteomics* 13, 2558-2572, doi:10.1074/mcp.M114.039156 (2014).
240 McDougald, D. *et al.* Defences against oxidative stress during starvation in bacteria. *Antonie Van Leeuwenhoek* 81, 3-13, doi:10.1023/a:1020540503200 (2002).
241 Raina, S., Missiakas, D. & Georgopoulos, C. The rpoE gene encoding the sigma E (sigma 24) heat shock sigma factor of Escherichia coli. *EMBO J* 14, 1043-1055, doi:10.1002/j.1460-2075.1995.tb07085.x (1995).
242 Xue, X., Tomasch, J., Sztajer, H. & Wagner-Dobler, I. The delta subunit of RNA polymerase, RpoE, is a global modulator of Streptococcus mutans environmental adaptation. *J Bacteriol* 192, 5081-5092, doi:10.1128/JB.00653-10 (2010).
243 Jin, J. *et al.* Mechanism analysis of acid tolerance response of bifidobacterium longum subsp. longum BBMN 68 by gene expression profile using RNA-sequencing. *PLoS One* 7, e50777, doi:10.1371/journal.pone.0050777 (2012).
244 Kleerebezem, M. *et al.* Lifestyle, metabolism and environmental adaptation in Lactococcus lactis. *FEMS Microbiol Rev* 44, 804-820, doi:10.1093/femsre/fuaa033 (2020).
245 Ercan, O., den Besten, H. M. W., Smid, E. J. & Kleerebezem, M. The growth-survival trade-off is hard-wired in the Lactococcus lactis gene regulation network. *Environ Microbiol Rep*, doi:10.1111/1758-2229.13073 (2022).
246 Sanchez, B., Noriega, L., Ruas-Madiedo, P., de los Reyes-Gavilan, C. G. & Margolles, A. Acquired resistance to bile increases fructose-6-phosphate phosphoketolase activity in Bifidobacterium. *FEMS Microbiol Lett* 235, 35-41, doi:10.1016/j.femsle.2004.04.009 (2004).
247 Helmann, J. D. The extracytoplasmic function (ECF) sigma factors. *Adv Microb Physiol* 46, 47-110, doi:10.1016/s0065-2911(02)46002-x (2002).
248 Hecker, M. & Volker, U. General stress response of Bacillus subtilis and other bacteria. *Adv Microb Physiol* 44, 35-91, doi:10.1016/s0065-2911(01)44011-2 (2001).
249 Kien, C. L. Digestion, absorption, and fermentation of carbohydrates in the newborn. *Clin Perinatol* 23, 211-228 (1996).
250 Chung, M. *et al.* FADU: A Feature Counting Tool for Prokaryotic RNA-Seq Analysis. *bioRxiv*, 337600, doi:10.1101/337600 (2018).
251 Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 12, 357-360, doi:10.1038/nmeth.3317 (2015).
252 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).
253 Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol* 11, R106, doi:10.1186/gb-2010-11-10-r106 (2010).
254 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139-140, doi:10.1093/bioinformatics/btp616 (2010).
255 Maere, S., Heymans, K. & Kuiper, M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* 21, 3448-3449, doi:10.1093/bioinformatics/bti551 (2005).
256 UniProt, C. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res* 49, D480-D489, doi:10.1093/nar/gkaa1100 (2021).
257 Achanta, P. S. *et al.* Quantum mechanical NMR full spin analysis in pharmaceutical identity testing and quality control. *Journal of pharmaceutical and biomedical analysis* 192, 113601, doi:10.1016/j.jpba.2020.113601 (2021).
258 Tang, Y. *et al.* Quantum Mechanics-Based Structure Analysis of Cyclic Monoterpene Glycosides from Rhodiola rosea. *Journal of Natural Products* 83, 1950-1959, doi:10.1021/acs.jnatprod.0c00212 (2020).
259 Mehta, P., Goyal, S. & Wingreen, N. S. A quantitative comparison of sRNA-based and protein-based gene regulation. *Mol Syst Biol* 4, 221, doi:10.1038/msb.2008.58 (2008).

260 Gottesman, S. *et al.* Small RNA regulators and the bacterial response to stress. *Cold Spring Harb Symp Quant Biol* 71, 1-11, doi:10.1101/sqb.2006.71.016 (2006).

261 Storz, G., Vogel, J. & Wassarman, K. M. Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell* 43, 880-891, doi:10.1016/j.molcel.2011.08.022 (2011).

262 Desnoyers, G., Bouchard, M. P. & Masse, E. New insights into small RNA-dependent translational regulation in prokaryotes. *Trends Genet* 29, 92-98, doi:10.1016/j.tig.2012.10.004 (2013).

263 Papenfort, K. & Vanderpool, C. K. Target activation by regulatory RNAs in bacteria. *FEMS Microbiol Rev* 39, 362-378, doi:10.1093/femsre/fuv016 (2015).

264 Sesto, N., Wurtzel, O., Archambaud, C., Sorek, R. & Cossart, P. The excludon: a new concept in bacterial antisense RNA-mediated gene regulation. *Nat Rev Microbiol* 11, 75-82, doi:10.1038/nrmicro2934 (2013).

265 Argaman, L. *et al.* Novel small RNA-encoding genes in the intergenic regions of Escherichia coli. *Curr Biol* 11, 941-950, doi:10.1016/s0960-9822(01)00270-6 (2001).

266 Kawano, M., Reynolds, A. A., Miranda-Rios, J. & Storz, G. Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in Escherichia coli. *Nucleic Acids Res* 33, 1040-1050, doi:10.1093/nar/gki256 (2005).

267 Leonard, S. *et al.* APERO: a genome-wide approach for identifying bacterial small RNAs from RNA-Seq data. *Nucleic Acids Res* 47, e88, doi:10.1093/nar/gkz485 (2019).

268 Umu, S. U. & Gardner, P. P. A comprehensive benchmark of RNA-RNA interaction prediction tools for all domains of life. *Bioinformatics* 33, 988-996, doi:10.1093/bioinformatics/btw728 (2017).

269 Eggenhofer, F., Tafer, H., Stadler, P. F. & Hofacker, I. L. RNApredator: fast accessibility-based prediction of sRNA targets. *Nucleic Acids Res* 39, W149-154, doi:10.1093/nar/gkr467 (2011).

270 Mann, M., Wright, P. R. & Backofen, R. IntaRNA 2.0: enhanced and customizable prediction of RNA-RNA interactions. *Nucleic Acids Res* 45, W435-W439, doi:10.1093/nar/gkx279 (2017).

271 Zhu, D. Q. *et al.* Genome-wide identification of small RNAs in Bifidobacterium animalis subsp. lactis KLDS 2.0603 and their regulation role in the adaption to gastrointestinal environment. *PLoS One* 10, e0117373, doi:10.1371/journal.pone.0117373 (2015).

272 Bottacini, F. *et al.* Global transcriptional landscape and promoter mapping of the gut commensal Bifidobacterium breve UCC2003. *BMC Genomics* 18, 991, doi:10.1186/s12864-017-4387-x (2017).

273 Duboux, S. *et al. The pleiotropic effects of carbohydrate-mediated growth rate modifications in B. longum NCC 2705* (Wageningen University and Research, 2022).

274 Fukunaga, T., Iwakiri, J., Ono, Y. & Hamada, M. LncRRIsearch: A Web Server for lncRNA-RNA Interaction Prediction Integrated With Tissue-Specific Expression and Subcellular Localization Data. *Front Genet* 10, 462, doi:10.3389/fgene.2019.00462 (2019).

275 Li, L. *et al.* BSRD: a repository for bacterial small regulatory RNA. *Nucleic Acids Res* 41, D233-238, doi:10.1093/nar/gks1264 (2013).

276 Oberg, T. S., Ward, R. E., Steele, J. L. & Broadbent, J. R. Transcriptome analysis of Bifidobacterium longum strains that show a differential response to hydrogen peroxide stress. *J Biotechnol* 212, 58-64, doi:10.1016/j.jbiotec.2015.06.405 (2015).

277 Averina, O. V., Zakharevich, N. V. & Danilenko, V. N. Identification and characterization of WhiB-like family proteins of the Bifidobacterium genus. *Anaerobe* 18, 421-429, doi:10.1016/j.anaerobe.2012.04.011 (2012).

278 Hung, M. N., Xia, Z., Hu, N. T. & Lee, B. H. Molecular and biochemical analysis of two beta-galactosidases from Bifidobacterium infantis HL96. *Appl Environ Microbiol* 67, 4256-4263, doi:10.1128/AEM.67.9.4256-4263.2001 (2001).

279 Papenfort, K. & Vogel, J. Small RNA functions in carbon metabolism and virulence of enteric pathogens. *Front Cell Infect Microbiol* 4, 91, doi:10.3389/fcimb.2014.00091 (2014).

280 Lee, C. M., Sieo, C. C., Wong, C. M. V. L., Abdullah, N. & Ho, Y. W. Sequence analysis of 16S rRNA gene and 16S–23S rRNA gene intergenic spacer region for differentiation of probioticsLactobacillus strains isolated from the gastrointestinal tract of chicken. *Annals of Microbiology* 58, 133-140, doi:10.1007/BF03179457 (2008).

281 Kwon, H. S. *et al.* Rapid identification of potentially probiotic Bifidobacterium species by multiplex PCR using species-specific primers based on the region extending from 16S rRNA through 23S rRNA. *FEMS Microbiol Lett* 250, 55-62, doi:10.1016/j.femsle.2005.06.041 (2005).

282 Orban, J. I. & Patterson, J. A. Modification of the phosphoketolase assay for rapid identification of bifidobacteria. *J Microbiol Methods* 40, 221-224, doi:10.1016/s0167-7012(00)00133-0 (2000).

283 Stothard, P., Grant, J. R. & Van Domselaar, G. Visualizing and comparing circular genomes using the CGView family of tools. *Brief Bioinform* 20, 1576-1582, doi:10.1093/bib/bbx081 (2019).

284 Scott, M. & Hwa, T. Bacterial growth laws and their applications. *Curr Opin Biotechnol* 22, 559-565, doi:10.1016/j.copbio.2011.04.014 (2011).

285 Korem Kohanim, Y. *et al.* A Bacterial Growth Law out of Steady State. *Cell Rep* 23, 2891-2900, doi:10.1016/j.celrep.2018.05.007 (2018).

286 Dunn, J. J. & Studier, F. W. T7 early RNAs and Escherichia coli ribosomal RNAs are cut from large precursor RNAs in vivo by ribonuclease 3. *Proc Natl Acad Sci U S A* 70, 3296-3300, doi:10.1073/pnas.70.12.3296 (1973).

287 Sulthana, S. & Deutscher, M. P. Multiple exoribonucleases catalyze maturation of the 3' terminus of 16S ribosomal RNA (rRNA). *J Biol Chem* 288, 12574-12579, doi:10.1074/jbc.C113.459172 (2013).

288 Jain, C. RNase AM, a 5' to 3' exonuclease, matures the 5' end of all three ribosomal RNAs in E. coli. *Nucleic Acids Res* 48, 5616-5623, doi:10.1093/nar/gkaa260 (2020).

289 Duboux, S., Muller, J. A., De Franceschi, F., Mercenier, A. & Kleerebezem, M. Using fluorescent promoter-reporters to study sugar utilization control in Bifidobacterium longum NCC 2705. *Sci Rep* 12, 10477, doi:10.1038/s41598-022-14638-4 (2022).

290 Pan, M., Nethery, M. A., Hidalgo-Cantabrana, C. & Barrangou, R. Comprehensive Mining and Characterization of CRISPR-Cas Systems in Bifidobacterium. *Microorganisms* 8, doi:10.3390/microorganisms8050720 (2020).

291 Han, X. *et al.* Characterization of CRISPR-Cas systems in Bifidobacterium breve. *Microb Genom* 8, doi:10.1099/mgen.0.000812 (2022).

292 Mariaule, V. *et al.* Digestive Inflammation: Role of Proteolytic Dysregulation. *Int J Mol Sci* 22, doi:10.3390/ijms22062817 (2021).

293 Cenac, N. *et al.* Role for protease activity in visceral pain in irritable bowel syndrome. *J Clin Invest* 117, 636-647, doi:10.1172/JCI29255 (2007).

294 Schmid, M. *et al.* Attenuated induction of epithelial and leukocyte serine antiproteases elafin and secretory leukocyte protease inhibitor in Crohn's disease. *J Leukoc Biol* 81, 907-915, doi:10.1189/jlb.0906581 (2007).

295 Galipeau, H. J. *et al.* Novel role of the serine protease inhibitor elafin in gluten-related disorders. *Am J Gastroenterol* 109, 748-756, doi:10.1038/ajg.2014.48 (2014).

296 Van Spaendonk, H. *et al.* Regulation of intestinal permeability: The role of proteases. *World J Gastroenterol* 23, 2106-2123, doi:10.3748/wjg.v23.i12.2106 (2017).

297 Vergnolle, N. Protease inhibition as new therapeutic strategy for GI diseases. *Gut* 65, 1215-1224, doi:10.1136/gutjnl-2015-309147 (2016).

298 Vergnolle, N. Protease-activated receptors as drug targets in inflammation and pain. *Pharmacol Ther* 123, 292-309, doi:10.1016/j.pharmthera.2009.05.004 (2009).

299 Sanchez-Navarro, A., Gonzalez-Soria, I., Caldino-Bohn, R. & Bobadilla, N. A. An integrative view of serpins in health and disease: the contribution of SerpinA3. *Am J Physiol Cell Physiol* 320, C106-C118, doi:10.1152/ajpcell.00366.2020 (2021).

300 Motta, J. P. *et al.* Modifying the protease, antiprotease pattern by elafin overexpression protects mice from colitis. *Gastroenterology* 140, 1272-1282, doi:10.1053/j.gastro.2010.12.050 (2011).
301 Langhorst, J. *et al.* Noninvasive markers in the assessment of intestinal inflammation in inflammatory bowel diseases: performance of fecal lactoferrin, calprotectin, and PMN-elastase, CRP, and clinical indices. *Am J Gastroenterol* 103, 162-169, doi:10.1111/j.1572-0241.2007.01556.x (2008).
302 Dewar, D. H. *et al.* Celiac disease: management of persistent symptoms in patients on a gluten-free diet. *World J Gastroenterol* 18, 1348-1356, doi:10.3748/wjg.v18.i12.1348 (2012).
303 Guandalini, S. & Assiri, A. Celiac disease: a review. *JAMA Pediatr* 168, 272-278, doi:10.1001/jamapediatrics.2013.3858 (2014).
304 Mitsinikos, F. T. & Pietzak, M. Celiac Disease, Wheat Allergy, and Non-Celiac Gluten Sensitivity: What's the Difference? *Adolesc Med State Art Rev* 27, 19-33 (2016).
305 Balakireva, A. V. & Zamyatnin, A. A. Properties of Gluten Intolerance: Gluten Structure, Evolution, Pathogenicity and Detoxification Capabilities. *Nutrients* 8, doi:10.3390/nu8100644 (2016).
306 Schumann, M., Siegmund, B., Schulzke, J. D. & Fromm, M. Celiac Disease: Role of the Epithelial Barrier. *Cell Mol Gastroenterol Hepatol* 3, 150-162, doi:10.1016/j.jcmgh.2016.12.006 (2017).
307 Gibson, G. R. *et al.* Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. *Nat Rev Gastroenterol Hepatol* 14, 491-502, doi:10.1038/nrgastro.2017.75 (2017).
308 James, K., Motherway, M. O., Bottacini, F. & van Sinderen, D. Bifidobacterium breve UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways. *Sci Rep* 6, 38560, doi:10.1038/srep38560 (2016).
309 Ramirez-Farias, C. *et al.* Effect of inulin on the human gut microbiota: stimulation of Bifidobacterium adolescentis and Faecalibacterium prausnitzii. *Br J Nutr* 101, 541-550, doi:10.1017/S0007114508019880 (2009).
310 Swanson, K. S. *et al.* The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of synbiotics. *Nat Rev Gastroenterol Hepatol* 17, 687-701, doi:10.1038/s41575-020-0344-2 (2020).

Summary

Probiotics are defined as “live micro-organisms that, when administered in adequate amounts, confer a health benefit to the host”. This definition implies that the health benefits attributed to the studied probiotic have to be demonstrated by *in vivo* trials conducted in the target population. There is an abundant list of positive clinical results associated with the administration of strains belonging to the *Lactobacillaceae* family or *Bifidobacterium* genus, but only few strains have been demonstrated to provide a consistent health effect in a congruent set of clinical trials. However, research on probiotic candidates belonging to the *Lactobacillaceae* family has generated clues suggesting that variations in manufacturing conditions might play a role in the discrepant outcomes of repeated trials, since they were proposed to modify the presence and activity of important probiotic effector molecules. Similarly, clinical trials conducted with bifidobacteria have led to controversial results, but the potential impact of manufacturing conditions has not been studied at the level of their functional molecular make-up to date. In **Chapter 2**, we review what is known on that subject, focusing on the potential effect manufacturing might have on the expression or presence of bifidobacterial niche factors and effector molecules. We highlighted the importance of growth conditions, and particularly the selection of the carbohydrate used as substrate as this was reported to influence the expression levels of several key bifidobacterial effector molecules (e.g., sortase-dependent pili, exopolysaccharides [EPS]).

The other chapters of the thesis focus on *B. longum* NCC 2705 and the serine protease inhibitor (serpin) it produces, using them as a model to study how production processes may influence i) the amount of this important bifidobacterial effector molecule; ii) the metabolism and overall physiology of the strain. In **chapter 3**, we investigate the carbohydrate-mediated regulation of serpin production in *B. longum* NCC 2705. We demonstrated that the production of serpin in *B. longum* is differently regulated depending on the carbon substrate utilized. Notably, the presence of glucose in the growth medium inhibited the production of the serpin protein, while fructose and galactose (and related oligosaccharides) facilitated its induction. These results led us to propose that serpin production in *B. longum* involves a mechanism that resembles the well described catabolite repression.

Chapter 4 deciphers the role of different pH regulation strategies during biomass production in a scaled-down fermentation process. The results described in this chapter established that depending on the selected pH regulation strategy, the levels of glucose consumed by *B. longum* NCC 2705 differ. Particularly, controlling the pH at 6 throughout the entire process led to glucose depletion before the end of the exponential phase, which enabled the production of serpin. Our data demonstrate that depending on the chosen pH control regime, glucose catabolite repression may be relieved, thus leading to high levels of serpin, which was further accompanied by an increased cellular membrane integrity and resistance to low pH.

In **chapter 5**, we studied the activity of specific promoters controlling carbohydrate utilization operons when the strain is grown on different single or mixed sugars. By implementing a fluorescent promoter reporter system and coupling it to flow cytometry, we observed that *B. longum* NCC 2705 behaved differently depending on the sugars it is exposed to. When the strain was grown on a mix of glucose and galactose, it displayed a standard diauxic behavior, glucose inhibiting the import and metabolization of galactose. In contrast, when exposed to a mix of glucose and arabinose the strain co-consumed the two sugars, suggesting that arabinose utilization is only partially subjected to the catabolite repression elicited by glucose. Overall, we confirmed the existence of glucose-mediated catabolite repression in NCC 2705, but also demonstrated that the level of repression elicited by the presence of glucose may be incomplete for certain alternative sugars (e.g., arabinose).

The results described in chapters 3, 4, and 5 clearly established the existence of glucose-mediated catabolite repression in *B. longum* NCC 2705 and its implication in the control of serpin production. In **chapter 6**, we further aimed at clarifying if glucose-catabolite repression might influence other physiological properties of *B. longum* NCC 2705. To this end, we used a combination of transcriptome analyses and physiological *in vitro* assays to assess the changes induced by growth on glucose or galactose as sole carbon source. We noticed two highly distinct phenotypes, which highlighted that depending on the carbon used during fermentation the physiology and metabolism of the strain was significantly modulated. Glucose grown NCC 2705 displayed a higher growth rate as compared to its galactose counterpart. This was reflected by several metabolic adaptations like the upregulation of nucleotide biosynthesis and ATP generation. In contrast, when cultured on galactose, the strain displayed a slow growth

rate, a reduced cell size and a strikingly expanded catabolic flexibility phenotype, which was supported by a variety of transcriptional changes related to carbon source import and central energy metabolism. In addition, galactose grown *B. longum* NCC 2705 exhibited an improved cell integrity and an enhanced acid-stress resistance as compared to glucose grown cells. To the best of our knowledge, this is the first demonstration of an association between reduced growth rate, expansion of catabolic flexibility, modified central carbon metabolism and increased pH stress resistance in bifidobacteria. Altogether, the results of **chapter 6** underline that not only growth outcomes but also industrially important physiological features of *B. longum* NCC 2705 are differently affected depending on the carbon source utilized.

In **chapter 7,** we deepened our understanding of the carbohydrate substrate-based regulation mechanisms in *B. longum* NCC 2705. Comparing glucose and galactose grown cells, we noted that serpin mRNA and protein levels did not correlate with each other, pointing to the involvement of a post-transcriptional regulatory mechanism. Combining transcriptome data and different RNA-RNA interaction prediction tools, we identified a *trans*-acting sRNA encoded in the conserved intergenic 16S-23S rRNA genomic region, that we named sBLO-161. This sRNA is predicted to efficiently bind by its 5' and 3' ends to the serpin (*BL0108*) mRNA and to form two internal stem-loops that could inhibit translation of the targeted mRNA. Further *in silico* analysis predicted sBLO-161 to form highly similar RNA-duplex structures with a range of mRNAs spread throughout the genome of NCC 2705. It suggested that this sRNA might control additional important functions in this strain, including stress response, cell wall synthesis, carbohydrate transport and metabolism. Finally, we show that sBLO-161 is relatively well conserved in different *B. longum* subspecies, although its regulatory role may be different in *B. longum* subsp. *infantis* ATCC 15697 as the 5' end of the sBLO-161 homologue is slightly modified in this strain. Overall, we describe the first bifidobacterial sRNA that is predicted to exert a global regulatory function in *B. longum* NCC 2705.

In chapter 8, we bring together the main findings of this thesis and further discuss potential future research directions. After revising our initial proposed serpin regulation model, we briefly recapitulate the evidence supporting the existence of a glucose-catabolite repression mechanism in bifidobacteria. We further propose the newly identified *trans*-acting sBLO-161 to act in concert with the previously identified Lac-I

type regulatory proteins AraQ and MalR1 to orchestrate catabolite-repression in NCC 2705. Then, we discuss how sBLO-161 may be produced (i.e., by splicing of the 16S-23S polycistronic rRNA) and regulated (i.e., by ribosomal RNA levels), and suggest that this newly discovered *trans*-acting sRNA regulates a wide range of growth rate associated functions in NCC 2705. Finally, based on the knowledge gathered in this thesis we propose the development of improved *B. longum* NCC 2705 preparations susceptible to locally deliver high levels of serpin for the management of two inflammatory diseases, i.e., celiac disease and IBD. We consider that the intestinal location of the disease must be taken in account when developing a tailor-made probiotic. Therefore, we suggest using serpin pre-loaded NCC 2705 in the case of celiac disease that principally affects the duodenum, and the design of a tailored synbiotic to ensure the production of serpin in the colon in the case of IBD.

In conclusion, the results obtained during this thesis underline that industrial probiotic applications may benefit from a deeper understanding of fundamentals aspects of bacterial physiology and its impact on effector molecule expression, to provide a path to the development of effect-oriented, robust, and reproducible probiotic concepts.

Acknowledgements / Remerciements

In English

I told many friends and colleagues that my aim was to finalize my PhD before my 40th birthday. I can now say that I have managed to do so, but I am not sure I fully realize the journey it has been. I am thankful to Nestlé, especially to **S. Palzer** and **R. Behringer** for giving me the possibility and support to pursue a PhD while remaining an employee involved in some key projects. When this opportunity was offered to me, I jumped at the chance without giving it too much thought. Looking back, I think it was the right approach. Those 4 years have turned out to be the challenge of a lifetime. Being able to manage studying, work, family, and everything else was a real balancing act, much like a tightrope walking exercise. I thought I was going to fall more than once. I was lucky to be surrounded by many people who have helped me to keep my balance and stay on the rope, and I would like here to thank them all.

First, I would like to thank my co-promotors from Nestlé. **Annick** and **Christoph,** you believed in me and made everything possible so that I could carry out my PhD in a set up that would not drastically impact the life of my family. Without the two of you, nothing would have been possible. **Annick**, you took care of me throughout the thesis, even after leaving Nestlé. You helped me on many occasions to find the necessary motivation to keep on progressing. Furthermore, your keen eye for detail certainly brought my thesis to another level. **Christoph, Jeroen** and **Biljana**, who succeeded as my group leaders and hence co-promotors during the thesis. The three of you, had to reorganize resources in the group and regularly justify the time I spent on my research. Without you, I would not have been able to spend enough time to complete and finalize my PhD before my 40th birthday, that is for sure.

I can not write acknowledgments without thanking **Michiel**, my main promotor. On top of being an amazing scientist, you also are a wonderful person. We first met during a break at the LAB symposium, thanks to **Annick**. You were part of the organizing committee (and therefore very busy) but took the time to talk with us. I still remember your reaction when we first proposed the topic. “That is not going to be an easy one” is what you said. I think that has proven to be true… But the care and the time you dedicated to me, all throughout the project, enabled me to go through the many hiccups and hurdles I faced.

I feel very honored to say that I was one of your students and will certainly keep that close to my heart for the rest of my life.

To all my (co-)promotors, **Christoph**, **Jeroen**, **Biljana**, **Annick** and **Michiel**, I will always remember the numerous laughs we had during our regular calls. I will especially remember the ones during the pandemic. They were a real breath of fresh air during those difficult times and motivated me to keep on progressing.

Furthermore, I would like to thank all my colleagues from Nestlé, who have contributed greatly in one way or another to my work. **Gaby**, your help to make my research relevant and visible was instrumental for the success of my thesis. **Mireille**, I learned a lot with you on how an ELISA should be properly developed and applied. Furthermore, I am lucky to be part of the **Technical Microbiology** team. As a PhD student, it is of the upmost importance to feel "at home" somewhere, and you (**Aliénor, Anne, Axel, Biljana, Claudia, Govind, Hannah, Josette, Lijuan, Nicole, Patrick, Rosa, Sarah**) all made that possible. I am proud to say that I am part of the "The microbes" (as per our official WhatsApp group header). A special thanks to **Anne**, **Josette**, **Nicole**, **Claudia** and **Aliénor** for your technical help on so many occasions, but also for maintaining a lab environment that is always functional. The value of that work may sometimes be underevaluated, but without it, there would simply be no great science.

To all the students I supervised, **Muzi**, **Sam**, **Bastien**, **Myrthe**, **Patrick** and **Cas**, thank you so much for your help throughout this scientific journey. **Muzi**, you are my first student who decided to pursue a PhD and your talents go way beyond science. I am really honored you accepted to design and draw the cover of my thesis; thank you so much. **Sam**, **Bastien**, **Patrick** and **Cas**, the four of you brought a lot of joy and positive energy to our group. I will especially remember **Sam**'s "joke of the week". **Cas**, you made me discover a new type of scientific literature, and as I write, I am finishing the book you gave to me. I have a special thought for you **Myrthe**. We could not pursue the internship in the way we had wished, as the pandemic struck. Packing your belongings and "throwing you" onto the last train back to the Netherlands so that you could be closer to your loved ones will remain a special moment in my life. To all my former students, you pushed me to think differently and to consider a different perspective on my data. I hence hope I transmitted to you as much as what you brought to me.

En français

Poursuivre un doctorat en cours d'emploi ne peut malheureusement se faire sans compromis. Dans les premières années de ma thèse, j'ai dû laisser de côté ma passion pour la musique et la direction d'ensemble de côté. Cette période m'a fait réaliser à quelle point la musique participait à mon équilibre personnel. Je remercie mes amis de **l'Union Instrumentale de Forel** qui m'ont donné l'opportunité de diriger à nouveau, après plusieurs années d'inactivité. Me retrouver à nouveau à la tête d'un ensemble m'a donné un surplus d'énergie qui m'a aidé à finaliser les dernières étapes de ma thèse.

Durant les quatre dernières années, je n'ai pas été physiquement et mentalement présent comme un frère, un fils ou un ami devrait l'être. A ma sœur **Aline**, ma mère **Hélène** et mon père **Jean-François**, merci du fond du cœur pour votre tolérance à mon égard durant cette période. Nous ne verbalisons pas toujours nos sentiments, mais je sais que vous étiez derrière moi ; les actes valent plus que les paroles. Merci également à mes beaux-parents **Heidi** et **Gilbert**, mes amis les plus proches et leurs familles, **Romain**, **Sébastien** et **Vincent,** ainsi qu'à tous mes proches qui feront le voyage aux Pays-Bas. Votre présence ce 9 décembre 2022 représente beaucoup pour moi.

Mes derniers remerciements, et non des moindres, vont aux deux femmes de ma vie ; mon épouse **Mélissa** et ma fille **Eloïse**. Sans aucun doute, vous avez été celles qui ont eu à vivre les côtés les plus éprouvants d'une thèse. L'ascenseur émotionnel qu'est un doctorat, combiné avec notre vie de tous les jours, a amené son lot de problème, d'incertitudes et de mauvaise humeur dans notre foyer. Et vous avez dû faire face à tout ceci. Mon âme sœur **Mélissa**, tu es une personne tellement passionnée, douce, qui prend soin des gens autour d'elle ; tu me soutiens à chaque fois que j'en ai besoin. Je n'aurai jamais assez de mots pour te remercier. **Eloïse**, tu avais 5 ans lorsque j'ai commencé. J'ai l'impression de ne pas avoir totalement réalisé à quel point tu as évolué ces quatre dernières années. Tu es aujourd'hui une jeune fille de 9 ans, pleine de joie, de talent, d'humour et d'empathie ; je suis tellement fier de toi. Vous êtes toutes les deux incroyables, et vous êtes les vraies stars de ce projet ; je vous aime.

About the author

Stéphane Duboux was born in Lausanne, Switzerland on 7th May 1983. After completing his secondary education (EPS Isabelle-de-Montolieu, Lausanne, Switzerland), he continued his studies in microbiology and biotechnology (EPSIC, Lausanne Switzerland) and obtained a "Federal Certificate of Capacity" (CFC) as well as a "Professional Maturity" Certificate in 2002. He then joined the Swiss National Research Cancer institute (ISREC, Lausanne; Prof. J. Huelsken's group) until 2007 where, as a laboratory assistant, he contributed to elucidate the role of the WNT pathway in skin cancer, resulting in his first peer-reviewed article published in "blood". In parallel, Stéphane studied music at the Conservatoire de Lausanne, Switzerland, obtaining certificates in trombone playing (2002), wind and brass band conducting (2007). He then joined Nestlé Research in 2007 to work in probiotic research, and in 2009 took over the coordination of the Nestlé Culture Collection (NCC). There, his first task was to standardize the collection so that it could be exploited for both research and industrial starter & probiotic production, leading to the FSSC 22'000 certification of NCC in 2011. Getting passionate about beneficial microbes, he gradually moved into a scientist role focusing on research and development involving the use of NCC strains. During the next 7 years, he was involved in or led several biotransformation projects within Nestlé Research, leveraging bacterial metabolism and genomics for the development of fermented foods and probiotics. From this period to the present day, he has published a list of 11 publications, including four first authored ones, and is co-inventor in 25 different published patent families. Until 2016, he was very active in the musical world, conducting several wind bands, including some at the highest level in Switzerland, and was invited to be an adjudicator at different contests. In 2018, Stéphane decided to further deepen his scientific knowledge and hence started his PhD research affiliated with Wageningen University and Research (The Netherlands), under the supervision of Prof. Dr. M. Kleerebezem, Dr. A. Mercenier, Dr. J. Muller and Dr. C. Bolten (later replaced by Dr. B. Bogicevic). His PhD research aimed at shedding light on the impact that manufacturing processes may have on the bioactivity of probiotics, with a specific focus on the model strain *Bifidobacterium longum* NCC 2705 and the serine protease inhibitor (serpin) effector molecule it produces. The results of his PhD research are presented in this thesis.

List of publications

Duboux S, Golliard M, Muller J A, Bergonzelli G, Bolten C J, Mercenier A, Kleerebezem M. Carbohydrate-controlled serine protease inhibitor (serpin) production in *Bifidobacterium longum* subsp. *longum*. Sci Rep 11, 7236, doi:10.1038/s41598-021-86740-y (2021). (this thesis)

Duboux S, Van Wijchen M, Kleerebezem M. The Possible Link Between Manufacturing and Probiotic Efficacy; a Molecular Point of View on *Bifidobacterium*. Front Microbiol 12, 812536, doi:10.3389/fmicb.2021.812536 (2021). (this thesis)

Duboux S, Muller J A, De Franceschi F, Mercenier A, Kleerebezem M. Using fluorescent promoter-reporters to study sugar utilization control in *Bifidobacterium longum* NCC 2705. Sci Rep 12, 10477, doi:10.1038/s41598-022-14638-4 (2022). (this thesis)

Jeannet G, Scheller M, Scarpellino L, Duboux S, Gardiol N, Back J, Kuttler F, Malanchi I, Birchmeier W, Leutz A, Huelsken J, Held W. Long-term, multilineage hematopoiesis occurs in the combined absence of beta-catenin and gamma-catenin. Blood 111, 142-149, doi:10.1182/blood-2007-07-102558 (2008).

Oliveira M, Bosco N, Perruisseau G, Nicolas J, Duboux S, Briand M, Blum S, Benyacoub J. *Lactobacillus paracasei* reduces intestinal inflammation in adoptive transfer mouse model of experimental colitis. Clin Dev Immunol 2011, 807483, doi:10.1155/2011/807483 (2011).

McCarville J L, Dong J, Caminero A, Bermudez-Brito M, Jury J, Murray J A, Duboux S, Steinmann M, Delley M, Tangyu M, Langella P, Mercenier A, Bergonzelli G, Verdu E F. A commensal *Bifidobacterium longum* strain improves gluten-related immunopathology in mice through expression of a serine protease inhibitor. Appl Environ Microbiol, doi:10.1128/AEM.01323-17 (2017).

Gangoiti J, van Leeuwen S S, Meng X, Duboux S, Vafiadi C, Pijning T, Dijkhuizen L. Mining novel starch-converting Glycoside Hydrolase 70 enzymes from the Nestle Culture Collection genome database: The *Lactobacillus reuteri* NCC 2613 GtfB. Sci Rep 7, 9947, doi:10.1038/s41598-017-07190-z (2017).

Gangoiti J, van Leeuwen S S, Gerwig G J, Duboux S, Vafiadi C, Pijning T, Dijkhuizen L. 4,3-alpha-Glucanotransferase, a novel reaction specificity in glycoside hydrolase family 70 and clan GH-H. Sci Rep 7, 39761, doi:10.1038/srep39761 (2017).

Duboux S, Ngom-Bru C, De Bruyn F, Bogicevic B. Phylogenetic, Functional and Safety Features of 1950s *B. infantis* Strains. Microorganisms 10, 203, doi:10.3389/fmicb.2021.812536, (2022).

Patel S H, Tan J P, Börner R A, Zhang S J, Priour S, Lima A, Ngom-Bru C, Cotter P D, Duboux S. A temporal view of the water kefir microbiota and flavour attributes. Innovative Food Science & Emerging Technologies 80, 103084, doi:10.1016/j.ifset.2022.103084 (2022).

Vatanen T, Ang Q Y, Siegwald L, …, Duboux S, …, Sakwinska O, Xavier R J. A distinct clade of *Bifidobacterium longum* in the gut of Bangladeshi children thrives during weaning. Cell, in Press (2022)

List of patents

Duboux S, Bergonzelli G, Mercenier A, Tangyu M. Serpin production, WO2019129808A1, 2019. (this thesis)

Duboux S, Duncan P, Golliard M, Kleerebezem M. Serpin production, WO2021001367A1, 2020. (this thesis)

Duboux S, Golliard M, Kleerebezem M. Serpin production, WO2021001368A1, 2020. (this thesis)

Alves T, Bergonzelli G, Duboux S, Duncan P, D'Urzo C, Verdu E. Methods and compositions using serpin producing bacteria, WO2021228805, 2021. (this thesis)

Arce Vera F, Bourqui B, Buetler T, Duboux S, …. Allaman I. 4-oxo-2-pentenoic acid and brain health, WO2013144077A1, 2013

Arce Vera F, Bourqui B, Buetler T, Duboux S, Foata F, Page N, Rezzi S. 4-oxo-2-pentenoic acid and cardiovascular health, WO2013144080A1, 2013

Arce Vera F, Bourqui B, Buetler T, Duboux S, Foata F, Page N, Rezzi S. 4-oxo-2-pentenoic acid and the health of the digestive tract, WO2013144083A1, 2013

Arce Vera F, Bourqui B, Buetler T, Duboux S, Guitard M, Foata F, Page N, Rezzi S. 4-oxo-2-pentenoic acid and skin pigmentation, WO2013144079A1, 2013

Arce Vera F, Bourqui B, Buetler T, Duboux S, Guitard M, Foata F, Page N, Rezzi S. 4-oxo-2 pentenoic acid and liver disorders, WO2013144085A1, 2013

Callanan J, Sakwinska O, Bourdin G, Duboux S, Bruttin A, Jankovic I, Ananta E. *Lactobacillus plantarum* NCC 2936 preparations and oral health, WO2014140080A1, 2014

Callanan J, Sakwinska O, Bourdin G, Duboux S, Ananta E. *Lactobacillus plantarum* CNCM I-4026 preparations and skin health, WO2016023688, 2015

Callanan J, Sakwinska O, Bourdin G, Duboux S, Bruttin A, Jankovic I, Ananta E. *Lactobacillus plantarum* NCC 2936 preparations and food preservation, WO2014140115, 2014

Bel-Rhlid R, Minehira Castelli K, Duboux S, Poquet L. Method of forming dihydroferulic acid, WO2016162227, 2016

Fleury Rey Y, Vafeiadi C, Duboux S, Ngom-Bru C, Mokdad C, Tastet C, Gysler C, Hansen C E. Flavour generation in food, WO2017198650A1, WO2017198652A2, 2017

Gangoiti J, van Leeuwent S, Dijkhuisen L, Vafeiadi C, Duboux S. Alpha glucans, WO2017207663A1, 2017

Gangoiti J, van Leeuwent S, Pijning J, Dijkhuisen L, Vafeiadi C, Duboux S. Branched alpha glucans, WO2018167032A1, 2018

Horcajada M-N, Duboux S, Poquet L. Bioconversion of Oleuropein, WO2019092068A1, WO2019092065, 2018

Duboux S, Horcajada M-N, Poquet L. Method of selecting a probiotic, WO2019092066A1, 2019

Benyacoub J, Duboux S, Prioult G, Bergonzelli G, Forbes-Blom E, Czarnecki G, Noti M. Probiotic combination for treatment of allergic disorders, WO2020127579A1, 2019

Benyacoub J, Duboux S, Prioult G, Bergonzelli G, Forbes-Blom E, Czarnecki G, Noti M. Probiotic combination for treatment of inflammatory-related Gastrointestinal disorders, WO2020070087, 2019

Radiguet S, Duboux S, Enault A. Flavoured fermented dairy product and method of preparation, WO2020128084, 2019

Schmitt C, Duboux S, Weiss J, Herz E, Oppen D… Dreher J. Vegan fermented pepperoni or salami analogue product, WO2022038209, 2021

De Bruyn F, Johnson K, Bogicevic B, Maes D, Duboux S. Mixture of HMOs and bifidobacteria, WO2022161865, 2022

Duboux S, Lamothe L, Aragao R, Rochat F. Beneficial effect of poaceae and/or leguminoseae fiber on survivability and metabolic activity of *Bifidobacterium longum,* WO2021233925, 2021

Duboux S, Lamothe L, Aragao R, Rochat F. Improvement of insoluble fibre fermentability and short chain fatty acid production by *Bifidobacterium longum,* WO2021233923, 2021

Completed training activities

Qualifying exams

Course: Cellular Signalling, EPFL, Switzerland, 2018
Course: Understanding statistics and experimental design, EPFL, Switzerland, 2018
Course: Advanced Food Microbiology, WUR, online, 2018

Discipline specific activities

Conference: Beneficial Microbes, Bastiaanse Communication, The Netherlands, 2018
Seminar: Gordon Research – AEM, GRC, USA, 2019
Conference: Gordon Research – AEM, GRC, USA, 2019
Course: Advanced Proteomics, VLAG, The Netherlands, 2019
Course: Intestinal Microbiome and Diet, VLAG, 2019
Symposium: Lactic Acid Bacteria Industrial Platform, LABIP, Italy, 2019
Course: Advanced Bioprocess Design, BioTech Delft, online, 2020
Symposium: Lactic Acid Bacteria Industrial platform, LABIP, online, 2021
Conference: Lactic Acid Bacteria Symposium, LAB13, online, 2021
Conference: European Fed. of Food Science and Technology, Elsevier, Switzerland, 2021

General courses

Presentation: Job forum, EPFL; Switzerland, 2018
Course: Scientific writing for biomedical articles, EPFL, Switzerland, 2019
Course: Responsible conduct in Biomedical Research, EPFL, Switzerland, 2019
Course: Scientific artwork, WGS, online, 2020
Course: Efficient Writing Strategies, WGS, online, 2020
Course: Bioinformatics with linux and python, WUR, online, 2021

Assisting in teaching and supervision activities

Sam de Jong (Delft University), internship supervision, 2018
Bastien Gentili (Marseille University), MSc supervision, 2019
Myrthe Van Wijchen (Wageningen University), internship supervision, 2020
Jan Patrick Tan (ETH Zurich), MSc supervision, 2021
Cas Zandbergen (Wageningen University), internship supervision, 2021

Other activities

Proposal writing, VLAG, 2018
Weekly lab meeting, Nestlé Research, 2018-2022
Literature peer-review, Scientific report, 2020
Literature peer-review, Microbial biotechnology, 2021
Literature peer-review, Applied and Environmental Microbiology, 2021

The work presented in this thesis was funded by Société des Produits Nestlé SA (Vevey, Switzerland) and was carried out at the Technical Microbiology research group, Nestlé Research (Lausanne, Switzerland). Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover design: Muzi Tangyu

Printed by: Digiforce | ProefschriftMaken